

Review

# Geospatial XAI: A Review

Cédric Roussel \*  and Klaus Böhm

i3mainz—Institute for Spatial Information and Surveying Technology, Mainz University of Applied Sciences,  
55128 Mainz, Germany

\* Correspondence: cedric.roussel@hs-mainz.de

**Abstract:** Explainable Artificial Intelligence (XAI) has the potential to open up black-box machine learning models. XAI can be used to optimize machine learning models, to search for scientific findings, or to improve the understandability of the AI system for the end users. Geospatial XAI refers to AI systems that apply XAI techniques to geospatial data. Geospatial data are associated with geographical locations or areas and can be displayed on maps. This paper provides an overview of the state-of-the-art in the field of geospatial XAI. A structured literature review is used to present and discuss the findings on the main objectives, the implemented machine learning models, and the used XAI techniques. The results show that research has focused either on using XAI in geospatial use cases to improve model quality or on scientific discovery. Geospatial XAI has been used less for improving understandability for end users. The used techniques to communicate the AI analysis results or AI findings to users show that there is still a gap between the used XAI technique and the appropriate visualization method in the case of geospatial data.

**Keywords:** explainable AI; geospatial ML; evaluation metrics; visualization



**Citation:** Roussel, C.; Böhm, K.  
Geospatial XAI: A Review. *ISPRS Int.  
J. Geo-Inf.* **2023**, *12*, 355. <https://doi.org/10.3390/ijgi12090355>

Academic Editor: Wolfgang Kainz

Received: 22 May 2023

Revised: 31 July 2023

Accepted: 29 August 2023

Published: 31 August 2023



**Copyright:** © 2023 by the authors.  
Licensee MDPI, Basel, Switzerland.  
This article is an open access article  
distributed under the terms and  
conditions of the Creative Commons  
Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Geospatial data include a wide range of information about the physical and natural world, such as location, topography, and weather, and have become increasingly important in many applications. Some examples of these applications include environmental monitoring, urban planning, and transportation management.

A promising approach to the challenges of analyzing and understanding Artificial Intelligence (AI) systems that has emerged in recent years is Explainable AI (XAI). XAI can be defined as “AI systems that can explain their rationale to a human user, characterize their strengths and weaknesses, and convey an understanding of how they will behave in the future.” [1]. In other words, the goal of XAI is to develop AI systems that can provide human-understandable explanations for their decisions and actions. XAI has the potential to enhance human capabilities and attributes such as comprehensibility and trust in various domains. A simple but concise definition of an explanation can be taken from Miller [2], who sees an explanation as “an answer to a why-question”. In XAI, the goal is to answer why questions from specific users, who may be developers, domain experts, or even laypeople. A term often used as a synonym for explainability is interpretability. However, we use interpretability as a different term. Some of the many research studies that use these two terms as distinct concepts are Gilpin et al. [3], Doran et al. [4], Broniatowski [5], and Bellucci et al. [6]. For example, Gilpin et al. [3] state: “We take the stance that interpretability alone is insufficient. In order for humans to trust black-box methods, we need explainability—models that are able to summarize the reasons for neural network behavior, gain the trust of users, or produce insights about the causes of their decisions. While interpretability is a substantial first step, these mechanisms need to also be complete, with the capacity to defend their actions, provide relevant responses to questions, and be audited. Although interpretability and explainability have

been used interchangeably, we argue there are important reasons to distinguish between them. Explainable models are interpretable by default, but the reverse is not always true”.

Our study will focus on geospatial XAI. Geospatial XAI, therefore, refers to systems that use XAI techniques to analyze and understand predictions and classifications with geospatial data and machine learning models. Geospatial data can be defined as: “information that describes objects, events or other features with a location on or near the surface of the earth. Geospatial data typically combine location information (usually coordinates on the earth) and attribute information (the characteristics of the object, event or phenomena concerned) with temporal information (the time or life span at which the location and attributes exist)” [7]. This means that the data are associated with geographical locations or areas and can be represented in the form of maps.

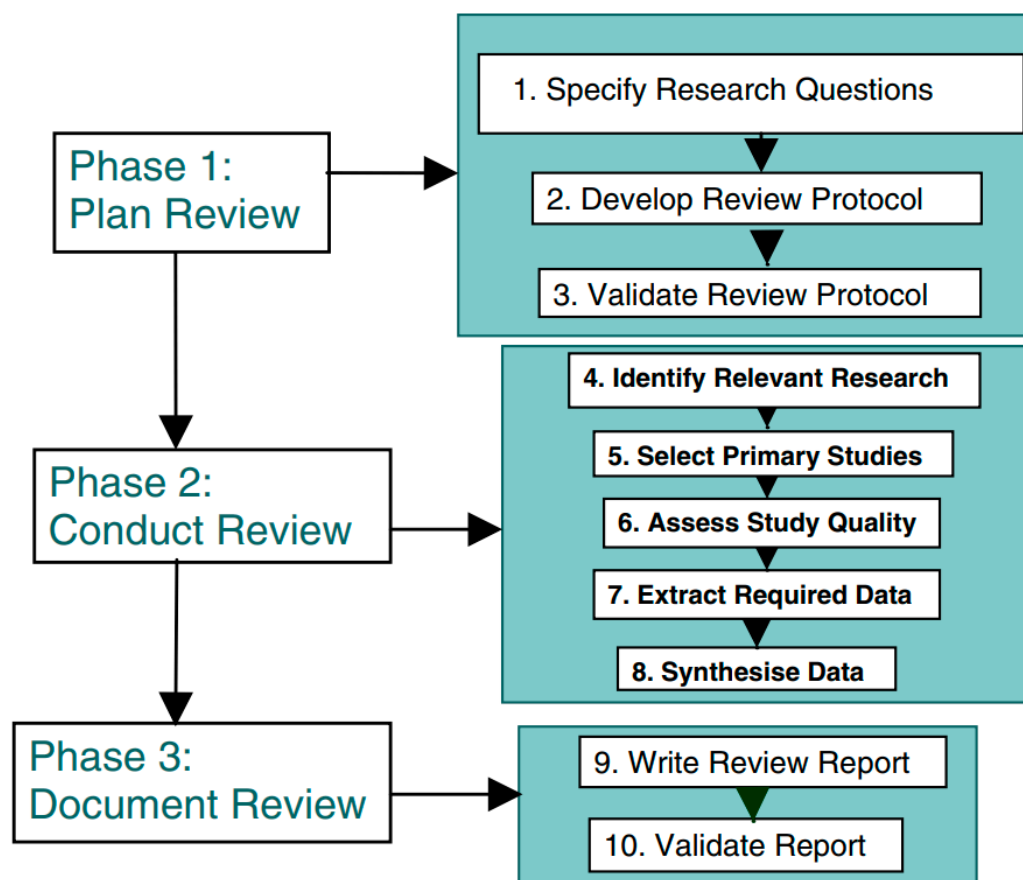
The importance of geospatial XAI rises from the decision-making process in domains including geospatial data. For example, in the case of natural disaster scenarios, AI systems must inevitably be transparent and understandable in the decision-making process to identify areas with a higher risk of flooding or landslides. In addition to statistical metrics for understandability, a suitable visualization is crucial for a better and faster understanding of geospatial data and machine learning model outputs. With the increasing complexity due to spatial and temporal data, more research on geospatial XAI is needed. Geospatial XAI requires different techniques and methods for an effective transparent and explainable output. This includes suitable visualizations of the geospatial data or XAI techniques, such as heat maps, plots, or even natural language explanations to show the reasoning behind the system’s decisions. By providing transparent and explainable output, geospatial XAI can support and guide stakeholders towards improved decision-making and more effective use of geospatial data.

In this study, a structured literature review is carried out to gather state-of-the-art information on geospatial XAI. This includes the main goal of research using geospatial XAI, the use cases, the applied machine learning models and XAI techniques, and the challenges and open issues for future work. The study will focus on explainability and not interpretability, as research sees those terms as different. In research, there are already many high-quality reviews of XAI in general. However, to the best of our knowledge, there are no XAI review studies that focus on use cases related to geospatial data. With our review, we provide a better insight into the explainability of machine learning models for geospatial data.

Section 2 presents the overall structured research methodology to identify relevant publications in this research field. In this step, five research questions are defined, which will be discussed later. The used database and the criteria for selecting relevant research are given. Section 3 will review the collected information through the defined research questions. The final section, Section 4, will provide a conclusion to this review.

## 2. Research Methodology

Systematic literature reviews (SLRs) are often used to collect, summarize, and review scientific topics in research. It is essential to use a structured approach to ensure that important research is not overlooked. This review will follow the guidelines of Brereton et al. [8], which are already widely used in state-of-the-art review articles, e.g., from Jenga et al. [9]. The review process of the guideline is shown in Figure 1.



**Figure 1.** Systematic literature review process [8].

The 10-stage review process is divided into 3 main phases, beginning with the planning of the review. This is where the research questions are defined and the research methodology (e.g., keyword strings, databases) is specified. In the second phase, the planned review is carried out in order to collect all the necessary research information for the final review. Research articles are collected from scientific databases, subjected to inclusion and exclusion criteria, and screened for relevance. Data from all relevant articles are then extracted and presented to provide an overview of XAI methods for geospatial data. The final phase will address the research questions from Step 1 and discuss the findings.

### 2.1. Research Questions

Research questions are defined for a structured review of the state-of-the-art, covering the most important information and achievements from the identified research. The following research questions will be investigated in this review:

1. What were the objectives of the paper?
2. Which ML algorithms were used?
3. Which XAI techniques were used, and which user group was addressed?
4. Was an evaluation carried out and if so, for what purpose?
5. What challenges or future work were identified in relation to XAI in geospatial data?

The focus is on question three, the used XAI techniques, and question five, the challenges and future work. The other questions are considered in order to better understand the overall research methodology, as XAI is based on the overall goal, the use case, and the machine learning model.

## 2.2. Database and Search Strategy

In the first path, MDPI open-access journals (<https://www.mdpi.com/>, accessed on 10 March 2023) are used to evaluate different search strings. After a rough relevance check of the results, the most suitable search terms are then used in Google Scholar (<https://scholar.google.com/>, accessed on 13 March 2023). Due to the novelty of the research field, all types of published available documents should be considered in this research, but the focus is on research papers from journals. Google Scholar covers many different databases and academic publishers ([https://scholar.google.com/citations?view\\_op=top\\_venues&hl=en](https://scholar.google.com/citations?view_op=top_venues&hl=en), accessed on 15 March 2023), including journals, conference proceedings, books, technical reports, university repositories, and more across the web. A few of the included academic publishers and open-access distribution services that are well known are: MDPI, IEEE (<https://www.ieee.org/>, accessed on 15 March 2023), Springer Link (<https://link.springer.com/>, accessed on 15 March 2023), Science Direct (<http://www.sciencedirect.com/>, accessed on 15 March 2023), and ArXiv (<https://arxiv.org/>, accessed on 15 March 2023).

As mentioned above, Google Scholar also includes non-peer-reviewed studies, which can be a criterion for not choosing this database. However, since geospatial XAI is a very young field of research and not all research is published in high-ranking journals, it is advantageous to collect all research in the first place and then screen them for quality. The first search in MDPI uses the following search strings: “GeoXAI”, “Geospatial XAI”, and “Spatial XAI”.

The most suitable string is used in the Google Scholar search, depending on the number of research articles returned.

## 2.3. Selection Criteria

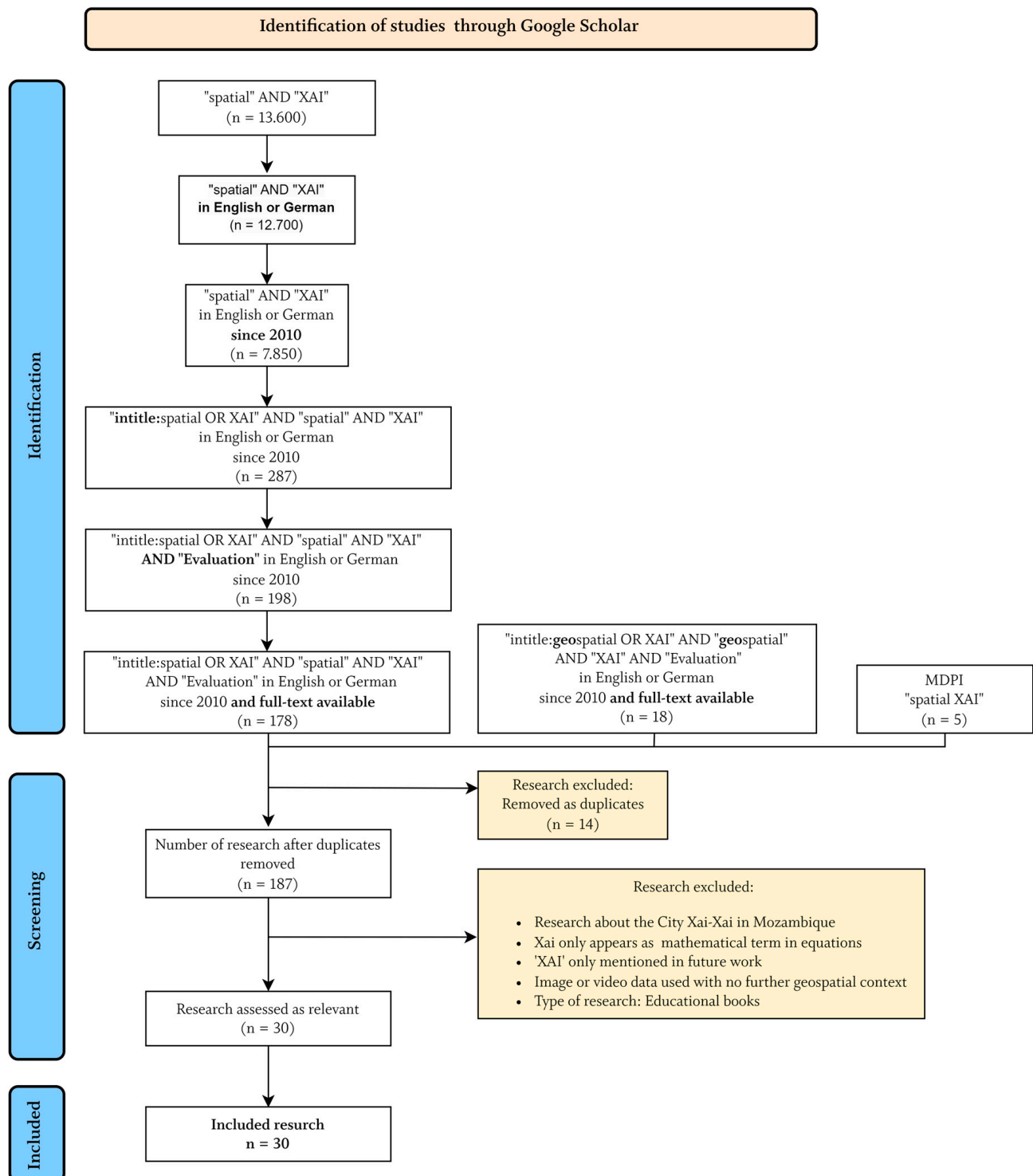
An article search in Google Scholar usually returns many results. In order to isolate the relevant studies, various inclusion and exclusion criteria are defined. The first important criterion is the year of publication. As XAI is a young field of research, all research published before the year 2010 is excluded. To illustrate the development and importance of this scientific field, the number of relevant studies for each year is visualized later. Secondly, all research that is not written in English or German is excluded. The last criterion is whether the full text of the article is available. In the next step, the results are screened for relevance through the title and abstract.

## 2.4. Collecting and Filtering

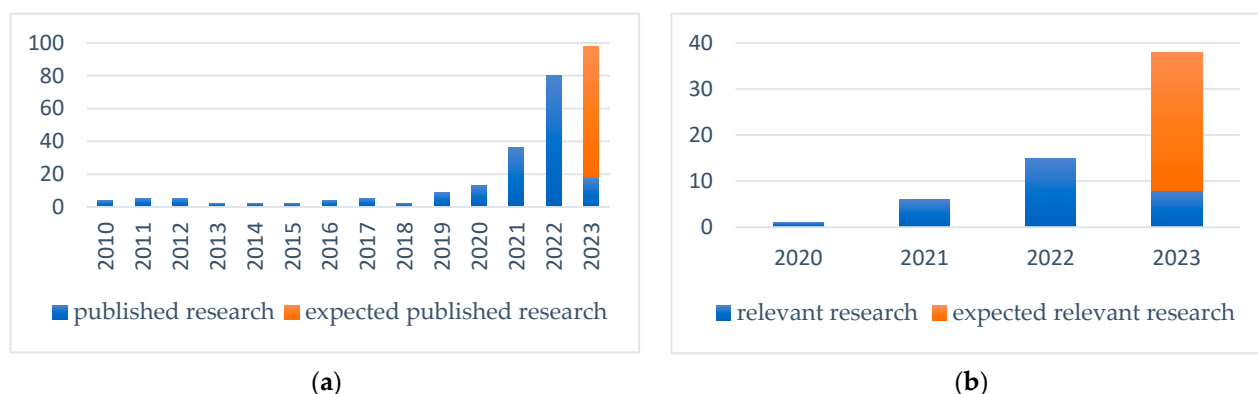
This chapter deals with the second phase of the systematic review process (Figure 1) by Brereton et al. [8]. The first step was to identify the suitable search terms. For the MDPI search, the strings “GeoXAI” and “Geospatial XAI” returned no results. The string “Spatial XAI” returned five articles. Although the word “spatial” can appear in many different use cases, such as mathematical theory or computer vision, it was the best choice to obtain results. Irrelevant articles can be discarded later. The Google Scholar search was conducted on 17 March 2023 and, together with 5 MDPI articles, resulted in 30 relevant studies. Figure 2 shows the specific number of results for each step of the search strategy. After the first iteration, a second search, with the addition of the term “geo,” was conducted using the same criteria.

Figure 3a shows the number of published research articles for each year after the full-text availability check and removal of all duplicates ( $n = 187$ ), and Figure 3b shows the number of studies included in this review ( $n = 30$ ). The orange color represents the estimated number of studies for the entire year 2023, upscaling the currently identified number of studies from 2023 to the full year of 365 days. Figure 3a clearly shows the increasing interest in XAI (relevance for geospatial XAI not considered at this time). Most of the research was published in the last two to three years. During the screening, research published before 2020 was not considered, as it was not relevant to this review. In the early years of this period, the term “XAI” only appeared in mathematical equations, or the research was about the city Xai-Xai in Mozambique. The first mention of XAI as Explainable

Artificial Intelligence in the identified research of this review was found in a presentation by Gunning in 2017 [10]. The first relevant research study was identified in the year 2020. Figure 3b also shows the increasing interest in this topic. Most of the research was published in the year 2022, and the upscaled expectation is even higher in 2023.

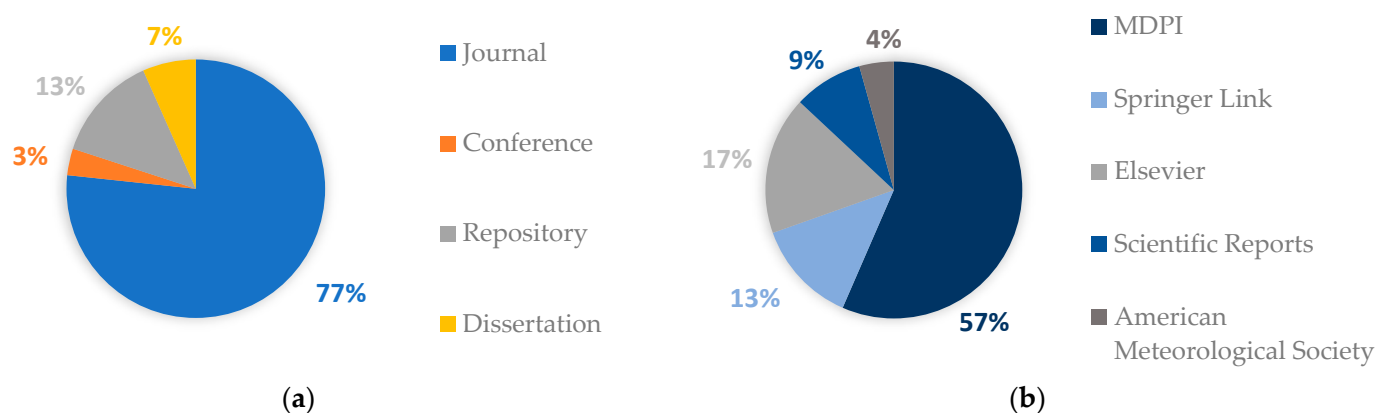


**Figure 2.** Search methodology applied to identify relevant studies.



**Figure 3.** (a) The number of published research per year after full-text availability check and removal of duplicates. (b) The number of included relevant research papers per year.

The relevant identified research in this review involves mainly journal articles. Other types of publications are two dissertations, one conference paper (Veran et al. [11] for the *IEEE International Conference on Big Data* (<https://bigdataieee.org/>, accessed on 21 March 2023)), or research published in scientific repositories (Hyper Articles en Ligne (<https://hal.science/>, accessed on 21 March 2023) ArXiv and Social Science Research Network (<https://www.ssrn.com/>, accessed on 21 March 2023)). Figure 4 shows the percentage distribution for the types and the journals separately. In the next chapter, the information on the identified research from the search strategy is presented and discussed through the research questions from Phase 1.



**Figure 4.** Distribution of selected studies per (a) publication type and (b) journal.

### 3. Present Findings

This chapter will present and discuss the current findings from Phase 2 of the structured literature review and discuss them before moving on to Phase 3. For this, the research questions are considered.

#### 3.1. RQ1—What Were the Objectives of the Paper?

The overall goals of the identified research are broad. The following table (Table 1) shows all the summarized use cases. Categories have been defined to cluster the identified research by use case. The categories are ordered by the number of matching research studies from highest to lowest.



**Table 1.** Categorized research into use cases—ordered by count.

Category	Count	References
Prediction of nature-related use cases	16	[12–27]
Uses cases in traffic and transport	5	[11,28–31]
Human-related use cases	5	[32–36]
Classification and scientific understanding in the medical domain	2	[37,38]
Potential location mapping	2	[39,40]

More than half of the research had the overall goal of prediction in nature-related use cases. This includes predicting slope failures/landslides [13,15,19,23] or avalanche hazards [20]; mapping natural disasters such as earthquakes [12,27] and wildfires [16,17]; or modeling other relationships with temperature [22,24,26] or pollution [18,21]. The next two major categories are research in traffic and transport and research in human-related use cases. For the use cases in traffic and transport, the goals were to predict car accidents [11], predict travel times [28] or demand for ride-sourcing [29], or tune and test machine learning models with travel data [30,31]. The category of human-related use cases covers a wide range of research, but it directly affects people’s lives. Graczyk-Kucharska et al. [35] analyze the career expectations of Generation Z candidates for human resources. Generation Z is defined as people born in the mid-to-late 1990s. For this, Graczyk-Kucharska et al. [35] use spatial data to explain spatiotemporal differentiations. Other research has focused on modeling urban expansion [33,36] or urban dynamics [32]. In another human-related study, in a use case other than urban modeling, Matuszelański and Kopczewska [34] try to model and understand customer churn using socio-demographic factors.

The first three categories cover two-thirds of the identified relevant research. Two research studies are classified in the medical domain, dealing with the influence of spatial factors on diseases. Ahmed et al. [37] try to understand spatial factors contributing to lung and bronchus cancer mortality rates, while Temenos et al. [38] identify factors influencing the spread of the COVID-19 virus. Two further studies in one category are based on location mapping to find potential sites for gold mineralization [39] and suitable wind and solar power plants [40].

Although three major categories can be identified, the research initiatives still differ in their research objectives.

### 3.2. RQ2—Which ML Algorithms Were Used?

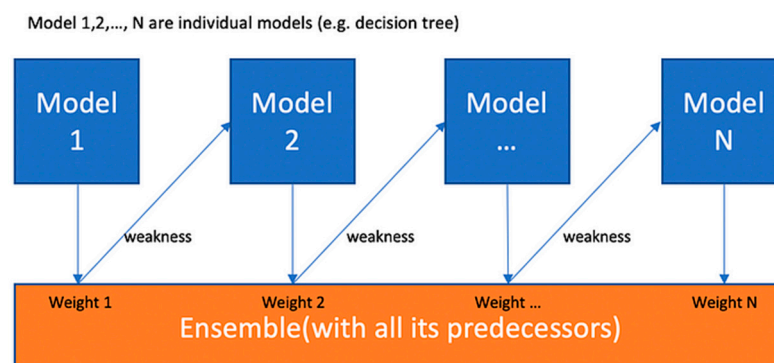
The second question looks at the used machine learning approaches and how the quality of these is assured. Before focusing on XAI techniques, a fitting machine learning model for the use case and main goal must be implemented. The benefit of a good model is seen later in the XAI component. If a model produces incorrect results, or predictions with too-high error rates, XAI becomes less powerful for the end user. In this case, XAI can be used to optimize the model, for example, by feature selection through feature importance. However, if XAI is used for scientific findings or end users, the machine learning model must provide adequate results, depending on the use case. The identified research implemented several different machine learning models. Very few studies used just one model, showing that it is not always clear which model to choose. Often, the best choice is to fit more than one model to the data and proceed with the most promising one. The following table (Table 2) shows the eight most commonly used approaches for machine learning models in the relevant research, along with the corresponding references.

**Table 2.** The eight most used machine learning approaches with references—ordered by appearance.

Machine Learning Model	Appearance	References
Boosting Approaches	19	Extreme GBM [11,20,24,28,29,31,33,34,36–38], Light GBM [20,25,38], Gradient BM [11,29], Explainable BM [13], Natural GB [20], Adaptive Boosting [20]
Neural Networks	17	Artificial ANN/Multilayer perceptron [12,18,22,26,27,29,35,38,40], Deep NN [14,28,37], Convolutional NN [21,26,39], Recurrent NN [32], Superposable NN [23]
Tree-based Models	15	Random Forest [12,13,15,16,18–20,27,29,37,38,40] Decision Tree [11,29,38]
Support Vector Machines	7	[13,15,19,28,29,38,40]
Linear Regression Models	5	LR [13,29], Linear Polynomial Regression [35], Lasso Regression [38], Generalized Linear Regression [37]
Logistic Regression	4	[17,19,23,34]
Clustering Algorithms	3	K Nearest Neighbor [13,18], DBSCAN [34]
Bayesian Approaches	2	Gaussian Process Regressor [38], Geostatistical Kriging [21]

The most commonly used machine learning models were boosting approaches, neural networks, and tree-based models. Although boosting approaches are also based on tree-based models, they are placed in a separate category due to the difference in the construction of these trees compared to single decision trees or random forests.

Boosting algorithms compute multiple tree-based models, such as decision trees. The same idea is used for random forest. In random forest, the computation of the different trees is based on bagging (bootstrap aggregation). This means that each tree is built with a random selection (with replacement) of instances in the training set [41]. Consequently, the trees are independent of each other. This is not the case for boosting algorithms. There, each tree is computed with the identified weaknesses/residuals of the previous tree (Figure 5). Together, they form an ensemble approach with the weights of each tree. There are different types of boosting algorithms. The most widely used kind in research is extreme gradient boosting [42], where trees are built in parallel rather than sequentially. Due to the high interest in boosting algorithms, research has focused on optimizing these algorithms, for example, by considering computational speed. A promising approach, also mentioned in Table 2, is the light GBM [43], where a leaf-wise (vertical) growth is used instead of a level-wise (horizontal) growth.

**Figure 5.** Boosting algorithm structure [44].

In addition to tree-based approaches, other models implemented in research are also quite popular. Models such as neural networks, SVM, and linear or logistic regression have been around for many years. Therefore, it is often advisable to implement one or more of these models to see how they perform on the data. Especially when new approaches are implemented, state-of-the-art benchmark models can be a good choice to compare the results with. This leads to an evaluation of the quality of the models. Besides many



different existing machine learning models, there are also numerous evaluation metrics for quality assessment. The following table (Table 3) shows all the used metrics in the research, with a brief description and the formula for their calculation.

**Table 3.** Used evaluation metrics for models—ordered by appearance. (TP = True Positives, TN = True Negatives, FP = False Positives, FN = False Negatives, P = all Positives, N = all Negatives).

Evaluation Metric	Appearance	Description	Formula
Overall Accuracy (OA) or R squared	20	The sum of correct classified individuals, divided by the total number of individuals.	$OA = \frac{TP + TN}{P + N}$ $R^2 = 1 - \frac{\sum_i (Actual_i - Predicted_i)^2}{\sum_i (Actual_i - \frac{1}{n} \sum_{k=1}^n Actual_k)^2} \quad (1)$
Root Mean Squared Error (RMSE)	12	The standard deviation of prediction errors.	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Actual_i - Predicted_i)^2} \quad (2)$
Recall	11	The ratio of correctly classified positive individuals out of the total of the class.	$Recall = \frac{TP}{TP + FN} \quad (3)$
The Area Under Receiver Operating Curve (AUROC)	10	Plot where the ROC shows the true positive rate (recall) on the y-axis against the false positive rate (specificity) on the x-axis. The area under the curve indicates the model performance.	
Mean Absolute Error (MAE)	10	Average of the difference between prediction and actual value.	$MAE = \frac{1}{n} \sum_{i=1}^n (Actual_i - Predicted_i) \quad (4)$
F1 score	10	The harmonic mean of precision and recall as accuracy measure.	$F1 \text{ score} = \frac{2 \times Recall \times Precision}{(Recall + Precision)} \quad (5)$
Precision	9	The ratio of correctly classified positive individuals to all positive individuals.	$Precision = \frac{TP}{TP + FP} \quad (6)$
Specificity	3	The ratio of correctly classified negative individuals out of the total of the class.	$Specificity = \frac{TN}{FP + TN} \quad (7)$
Kappa Coefficient	2	Ranges from −1 to 1; 0 indicates no better classification than a random classification.	$K = \frac{P_o + P_e}{1 - e} \quad (8)$
			where $P_e = \frac{TP + TN}{P + N} \quad (9)$
			and $P_e = \frac{P \times (TP + FN) + N \times (FP + TN)}{(P + N)^2} \quad (10)$
Negative Predictive Value (NPV)	1	The ratio of correctly classified negative individuals to all negative individuals.	$NPV = \frac{TN}{TN + FN} \quad (11)$
Mean Squared Error (MSE)	1	Average of the squares of the errors.	$MSE = \frac{1}{n} \sum_{i=1}^n (Actual_i - Predicted_i)^2 \quad (12)$
Mean Absolute Percentage Error (MAPE)	1	Average percentage of the difference between prediction and actual value.	$MAPE = \frac{1}{n} \sum_{i=1}^n \left  \frac{Actual_i - Predicted_i}{Actual_i} \right  \quad (13)$
Matthews Correlation Coefficient (MCC)	1	The correlation coefficient between actual individual and predicted individual.	$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{P \times (TP + FN) \times N \times (FP + TN)}} \quad (14)$

The most commonly used metric is the overall accuracy, using confusion matrix values for classification tasks or the R-squared value for regression tasks. These values always provide a good first overall impression of the model quality. However, they should never be used alone. Biased data can lead to a misinterpretation of these values. For this reason, research does not rely on one quality metric alone. Which metric to use depends on the use case and the data, starting with whether the problem is classification or regression. For example, the RMSE is a performance measure that cannot be reasonably used for classification tasks.

### 3.3. RQ3—Which XAI Techniques Were Used, and Which User Group Was Addressed?

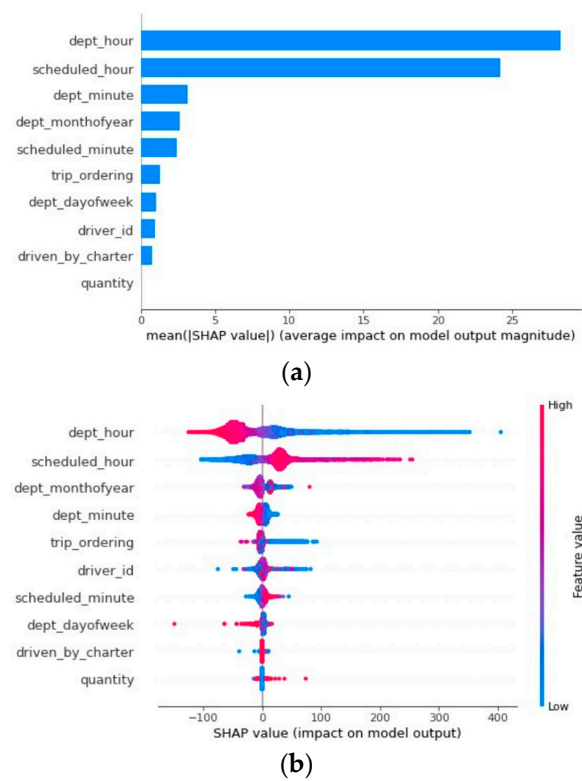
XAI can be used to achieve different goals. Research showed three major targets when using XAI: Explain to improve (23.3% of all research); explain to discover (76.7% of all research); and explain to control (20.0% of all research). Some research studies included several strategies for using XAI. The first one is to use XAI as a method to optimize the implemented machine learning model. For example, if the data contain many features, the XAI feature importance can be used to select more appropriate features to improve model quality and reduce complexity. The second option is the implementation of XAI for scientific findings, thereby limiting the understandability of XAI to researchers. There are no direct end users who need to understand the behavior of the model. The goal is to extract knowledge from data using machine learning and XAI. The third goal of XAI in research is to use the explanations to improve understandability for end users. There, it is inevitable to respect the knowledge of the end-user group in order to choose suitable visualization methods. Ideally, a user evaluation will be carried out.

This review will focus on the two XAI targets: explain to discover and explain to control. The following four chapters will present the most commonly used XAI techniques in research for these goals, starting with Shapley additive explanations (SHAPs) and local interpretable model-agnostic explanations (LIMEs). After that, visualization techniques combined with maps and other less-used but relevant techniques are presented.

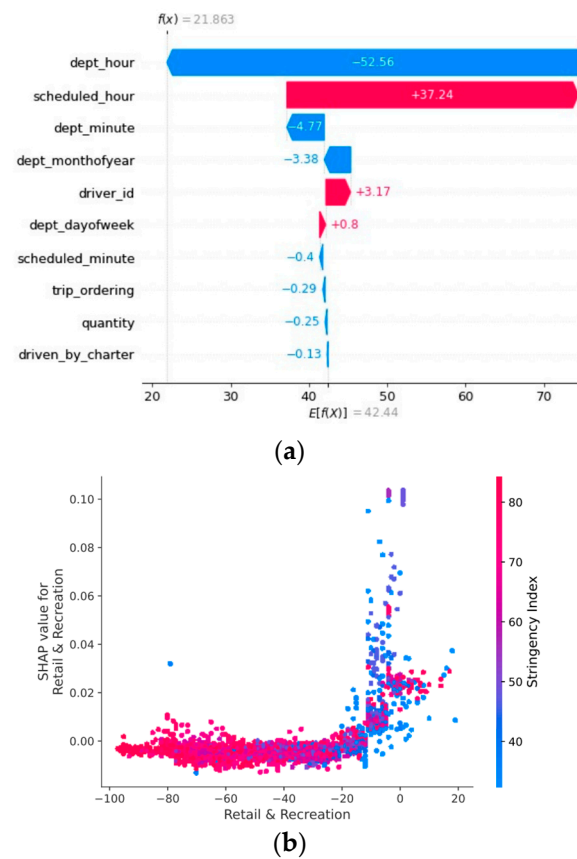
#### 3.3.1. SHAPs

By far the most widely used XAI technique was Shapley additive explanations [45]. SHAPs can provide local and global (through aggregations of Shapley values) explanations and are model-agnostic, meaning that they can be applied to any trained machine learning model. The goal of this technique is to compute the feature contributions for each particular prediction, to obtain a “unified measure of feature importance” [45]. The results of SHAPs can be visualized in different plots for different explanatory goals. The simplest is to aggregate all Shapley values to obtain a global explanation and present the mean contribution of all features in a bar plot (Figure 6a), as in the research of Ahmed et al. [28]. This method can be used in the same way as regression coefficients (feature importances) and gives a good first impression of which features are the most important. The disadvantage of this method is that it does not allow for negative contributions. For a more specific understanding of the contributions, other plots for local explanations are used. Four common plots for this are the SHAP summary plot or beeswarm plot (Figure 6b); the waterfall plot (Figure 7a), also known as the break-down plot or contribution plot; the dependency plot (Figure 7b); and the force plot (Figure 8).

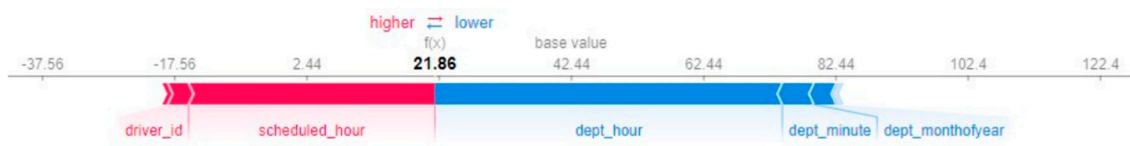
In the summary plot, each point represents a Shapley value of a feature of an instance in the data. The color of the point represents the value of the feature, and the position on the  $x$ -axis represents the contribution to the model output. Multiple points with similar contributions are jittered along the  $y$ -axis to provide a better view of the distribution. In the example of Ahmed et al. [28], the features `dept_hour` and `scheduled_hour` have the highest and the widest range of impact on the model output (Figure 6b), which is reflected in the mean contribution (Figure 6a). The coloring shows that, for example, the feature `dept_hour` has a more positive impact at lower feature values (blue color) and a negative but less impactful contribution at higher feature values (red color). These first two plots are for global explanations. Local explanations with SHAPs can be visualized with the waterfall plot (Figure 7a) or the force plot (Figure 8). Both show the contribution of the features for a particular model output, with the blue color showing features with a negative contribution and the red color showing features with a positive contribution. In these two examples of Ahmed et al. [28], the feature `dept_hour` has the highest negative impact, and `scheduled_hour` has the highest positive impact. The visualization starts with a base value, which can be the mean prediction, in this case,  $E[f(x)] = 42.44$ . Then, each feature pushes the value up or down with its contribution, ending with the final output ( $f(x) = 21.86$ ).



**Figure 6.** Examples of visualization plots of SHAP results in research for (a) the mean feature contribution [28] and (b) the summary plot or beeswarm plot [28].



**Figure 7.** Examples of visualization plots of SHAP results in research for (a) the waterfall plot, break-down plot, or contribution plot [28] and (b) the dependency plot [38].



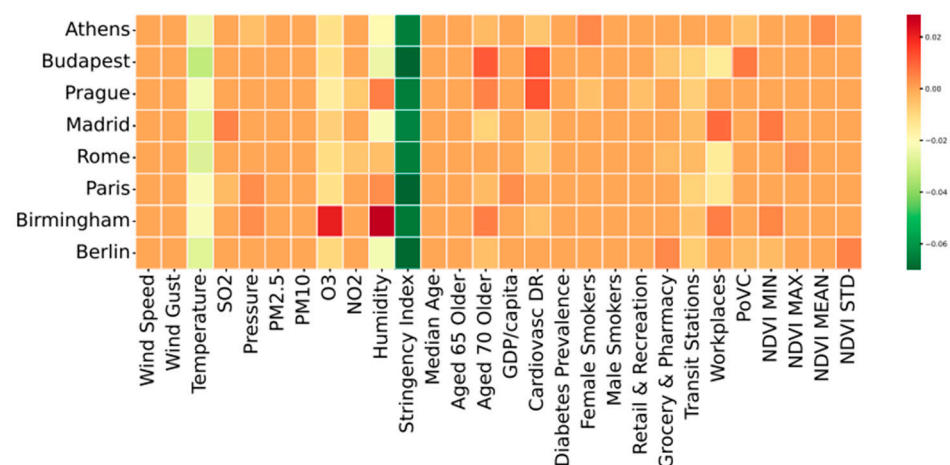
**Figure 8.** Example of visualization plot of SHAP results in research for the force plot [28].

A more detailed visualization between feature values and contributions can be achieved with the SHAP dependency plots, which can complement the summary plot. As in the summary plot, each point represents a data instance. The  $x$ -axis represents the feature values, and the  $y$ -axis represents the contribution through the Shapley values. In addition, custom coloring can be implemented for a more detailed visualization of interactions. For example, Temenos et al. [38] colorized their dependency plots regarding the Government Response Stringency Index (Figure 7b).

The five shown plots are the most commonly used ones, but further visualizations or slight modifications of these plots have been implemented in other research studies. For example, an aggregation of summary plots into a matrix was used to display a so-called interaction plot by Jena et al. [27]. Li [31] added bootstrap confidence to their dependency plots, which can show uncertainties in the XAI technique. Some studies did not use any visualization at all and presented their results in a table [18].

### 3.3.2. Local Interpretable Model-Agnostic Explanations

Another common XAI technique for local explanation is LIME [46], which is also model-agnostic and can be applied to different machine learning models. The goal of LIME is to use a local linear model to understand predictions, using perturbations. By perturbing the input of data samples, the predictions change, which can locally explain the prediction with positive and negative contributions from features. Temenos et al. [38] used LIME in their spatial epidemiology research for several cities. The results are presented in a matrix (Figure 9). As in the summary plot for SHAPs, the color of a parcel represents its influence on the model output, which can be positive or negative. This matrix, which is a kind of heat map, can be used to identify hotspots. In this example, we can see that the feature Stringency Index has a negative contribution in all cities.



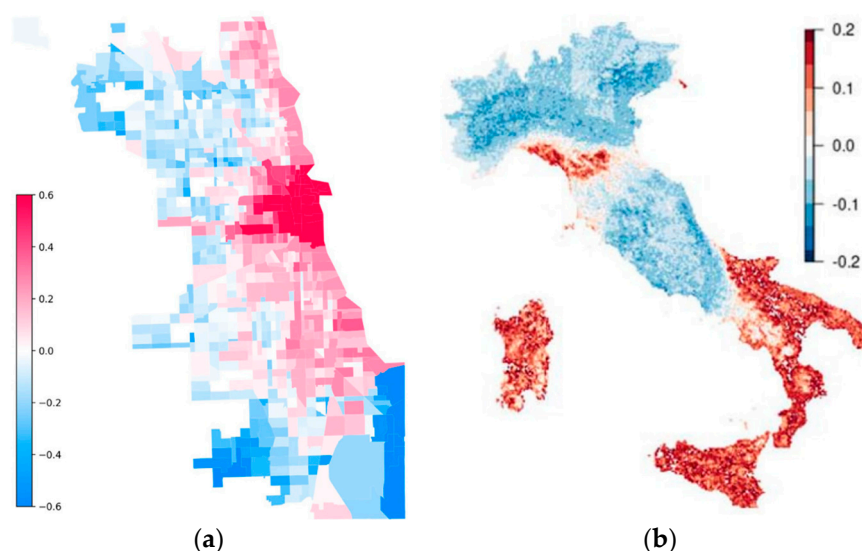
**Figure 9.** Example of local feature understanding with LIME [38].

Other possible plots are similar to the SHAP plots. For example, the waterfall plot can be used to show the contributions of the features to the final output, starting from a baseline such as zero [28]. As with SHAPs, adaptations of the framework are possible. For example, Jin [32] extended the LIME framework with a geographical context for Geographical LIME (GLIME). This enhances the spatial interpretability for the use case with geographical units such as states and counties.

### 3.3.3. Visualization on a Map

Even though the research studies involved geospatial data, not all of them visualized their results on a map. Some examples of visualizations are given here.

Li [31] showed one possible way of combining SHAP results with a map. In their SHAP summary plot, two of their features were the x and y coordinates of the data point. With the visualization in the summary plot, it is complex to understand higher and lower feature values in coordinates. For a better understanding, they displayed the impact of the coordinates separately on a map (Figure 10a), using the same commonly used coloring—from blue (negative impact) to red (positive impact). This makes it easier to understand which data points in which geographical locations have a negative or positive impact on the model output. A similar approach was used by Cilli et al. [16], who mapped the SHAP values of the Fire Weather Index in a wildfire occurrence prediction (Figure 10b). This technique allows for the identification of hot spots, which is not possible with the summary plot, splitting the coordinate pairs into two attributes. With a similar colorization, the mapping of the SHAP values can be a good contribution to a better understanding. Plotting local explanations next to a map and marking the considered area can also achieve better explainability [12].



**Figure 10.** (a) The spatial effect of location on the use case of ride-hailing demand, using the SHAP values [31]. (b) Mapped SHAP values of the feature Fire Weather Index in the prediction of wildfire occurrences [16].

Another way of including a map is to show non-influential data or untrustworthy predictions [18]. Together with, e.g., feature importance, untrustworthy results could be understood. A similar approach is to use two maps with different data inputs for the predictions due to natural variation, feature selection, and more [24,27]. This way, a counterfactual explanation could be achieved, where changes in the prediction due to data variance or the benefit of feature selection through SHAP values can be seen. Especially for non-experts, a visualization on a map can lead to a higher interest and understanding. For example, Liang et al. [21] used two maps to visualize the true feature values in the data and the predicted feature values, colored with the same quantiles within their range. This makes it possible to see a correlation between the two features more easily. The visualization can increase the trust of laypersons in the result, as they can see for themselves through the colors that both maps look somehow the same. A single correlation value is more accurate but less understandable, which may lead to a lack of trust among laypersons.

With all these possibilities, care must be taken not to use too many visualization components due to the risk of cognitive overload. Simple combinations of components on a map can easily lead to cognitive overload, as can be seen in Stadtler et al. [18].



Stadtler et al. [18] presented a map with simple, but multiple, colorizations and shapes, where the components overlap each other in some areas.

With less data, more differentiation in the visualization is possible. However, when the amount of data reaches a point where visualizations overlap each other and interpretation becomes difficult, splitting the results into different maps or plots should be considered.

### 3.3.4. Further XAI Techniques and Visualizations

The first three subsections of XAI techniques covered most of the research. Some other techniques have been applied for specific cases or applications. One XAI technique similar to using SHAP values is the feature importance. There are two ways of calculating feature importance. One would be to use linear models and take the regression coefficients [17]. The other one is to use permutation [16,21,37]. Permutations in the data change the prediction of the model. If a feature is changed, and the model output does not show major changes, then the feature is not contributing much to the prediction. If the output becomes less accurate, then the feature is more important to the model. The greater the increase in prediction error, the greater the permutation-based feature importance. The importance of all features can then be visualized in a bar plot like the mean SHAP values, in scatter plots [13], or simply in a table [18,21]. The disadvantage of the second way of calculating feature importance is that no negative importance is calculated. A higher permutation-based feature importance does not indicate whether the feature has a high positive or negative impact on the output. Regression coefficients, on the other hand, can have negative importance, like SHAP values. For random forest models, there is also the so-called Gini importance, which is calculated as the sum of the number of splits in all trees including the feature. This technique has been used by Sachit et al. [40] and visualized in a pie chart. This means that the sum of all values is one (or 100%) and the values can only be positive, similar to the pie chart of Youssef et al. [23].

Further XAI techniques were used, but they were not identified as relevant because they are less widely used. As we will not present each of them, we encourage exploring the following used XAI techniques, which are provided with the references:

- Integrated gradients by Mamalakis et al. [22] and Bommer et al. [26].
- Swarm plot using the SHAP values by Pradhan et al. [39].
- Two-variable partial dependency plots by Ahmed et al. [37].
- Randomized Input Sampling Explanation by Zhang and Zhao [30].
- Relationship plot by Graczyk-Kucharska [35].

### 3.4. RQ4—Was an Evaluation Carried Out and If So, for What Purpose?

The results of research question 4 show that most of the time, XAI is used to make scientific discoveries in the data. In this case, an evaluation is not always necessary, as no one other than the researcher is involved. Similarly, when XAI is used to improve the quality of the model, an evaluation of understandability is rarely important. In this case, XAI is evaluated by the quality improvement of the machine learning model using the metrics mentioned in research question 2.

For the third way of using XAI—to improve user understandability—it is very important to evaluate the used XAI techniques. An evaluation is inevitable before the machine learning model, together with the XAI component, is applied in the real world, where users need to understand what is happening.

Doshi-Velez and Kim [47] define three levels of experimentation to evaluate the effectiveness of XAI for user understandability: application-grounded, human-grounded, and functionally grounded. The application-grounded evaluation is carried out with end tasks in the real world and with domain experts, taking the role of end users. This level is the most difficult to use because domain experts need to be involved in the evaluation, and they are not easily available. Furthermore, this method is time-consuming, but the experiment can be shortened by using partial tasks. Despite the disadvantages, this evaluation method will provide the best possible results for measuring the understandability



of the XAI technique. This evaluation should be considered especially in use cases, where decisions with a high impact on people's lives are made. In a human-grounded evaluation, the experiment is also carried out with people. However, the tasks are simpler, and the test users are laypersons rather than domain experts. The advantage is that no experts are needed, which saves money and time. The obvious disadvantage is the lack of knowledge in the feedback. The final level is the functionally grounded evaluation, where a formal definition of interpretability is used to evaluate the XAI technique. This can be, for example, the depth of decision trees. This evaluation only makes sense if the XAI technique has already been evaluated in another study using a human-grounded evaluation [48].

Research question 3 showed that 6 of the 30 identified studies used XAI to improve understandability. Three other studies mentioned using XAI for end users in future work. Cilli et al. [16] state: "[...] XAI, capable of creating learned models and decisions that are understood and effectively trusted by end-users." However, despite the importance of end-user understandability, these six research studies [11,16,24,25,33,39] conducted only a functionally grounded evaluation of the effectiveness of the XAI techniques used. This means that they rely on other research that has evaluated the explainability of, for example, SHAP values and the corresponding plot type. Due to the variance of use cases and data, the understandability can change significantly. Especially for highly responsible decision-making, at least a human-grounded evaluation should be carried out before applying the model in the real world. To demonstrate an example of an application-grounded evaluation, we briefly present the work of Tosun et al. [49]. This research study appeared in the Google Scholar search but was not identified as relevant, as no geospatial data were used. However, this study shows a good example of how to integrate an evaluation. In their work, they developed an XAI platform in the field of computational pathology. They present a prototype for breast core biopsy to assist pathologists. They used breast cancer images with different machine learning models such as decision trees, support vector machines, or logistic regression, which showed the best results in terms of accuracy. As for the explainable component, they present the influential features with labels. In their prototype, they add a "why" button, which provides users with more explainable information about the AI system's decisions. In one of their evaluation studies, Tosun et al. [49] conducted experiments with pathologists, taking the role of domain experts. One of the evaluation metrics was to identify how much time the domain expert could save with the assistance of the prototype. Overall, their prototype-driven diagnostic process took 56% less time.

Out of 30 relevant research studies, 6 considered end-user evaluation, but no study conducted a human-level evaluation with domain experts. Before applying XAI prototypes in the real world, more evaluations should be considered to confirm the effectiveness of XAI. Otherwise, the possibility exists that the XAI technique and its visualization may not be suitable for the use case and may not help users to understand the decisions behind the AI system. It is highly recommended to include users in the process; this is also stated by a recent study by Xing and Sieber [50]: "[...] if they are not actively participating in the generation of explanations, in 'explaining the explanations,' then it becomes difficult to guarantee the effectiveness of XAI even if it is technically correct".

### 3.5. RQ5—What Challenges or Future Work Were Identified in Relation to XAI in Geospatial Data?

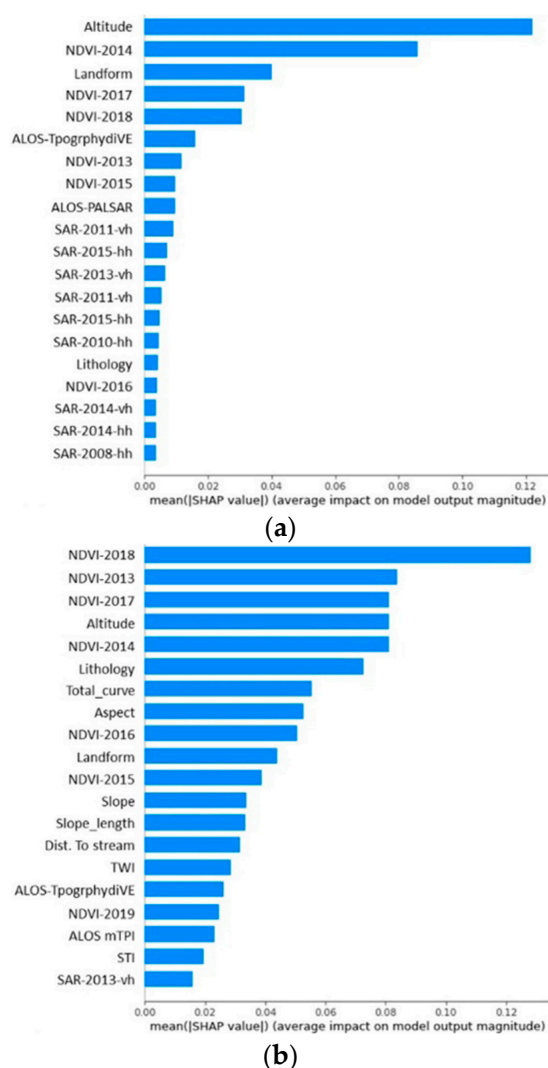
The major challenges and future work arising from the research studies were either general XAI issues or those specific to the additional geospatial component. This chapter presents some of the challenges and future work from the research studies. An overall view of the research gap for geospatial XAI is given later in Chapter 4: Conclusion.

Due to the diversity of use cases, data, machine learning models, and XAI techniques, implementing XAI can sometimes be challenging. Several research studies have identified general human-centered challenges. XAI can receive a high level of responsibility, as black boxes should not be given the power to make decisions that deeply affect human lives, as stated by a research study outside of the 30 identified studies [49]. One of the challenges

to be overcome in order for XAI to be used well, in line with its responsibilities, is to increase user confidence and satisfaction with explanations [25]. Maxwell et al. [13] found a challenge in that a model output with a large number of features is harder for users to understand. In order to determine the usefulness of XAI, so that users understand model outputs and decisions, more user evaluations need to be computed [19,28].

Possible future work regarding geospatial XAI was mentioned by Maxwell et al. [13] and Pradhan et al. [39], who stated that an interactive raster map to obtain local interpretations by clicking would be valuable to improve the understandability for end users. This would allow end users to explore for themselves what they want to understand about the AI system and about where they want to understand, in terms of the geospatial context. Other research studies have also shown that more robust and generalized explanations are needed [20] and that it is complex to compare XAI results in geoscience with the current state-of-the-art [22].

Considering SHAPs, several different challenges have been identified. For example, SHAPs were found to produce different results for different machine learning models [12,18] (Figure 11) but also different results in different geospatial study areas [12]. Al-Najjar et al. [15] mention that further additive properties of SHAPs need to be investigated. Similar future work is mentioned by Li [31] or Temenos et al. [38], who implement new SHAP methods such as fastSHAP for more efficiency.



**Figure 11.** Different mean SHAP values for (a) random forest and (b) support vector machine [15].

A similar challenge with the use of different machine learning models was identified by Li [31] in that models treated correlations differently, affecting interpretation. Other research studies have also reported issues with correlations. Maxwell et al. [13] state that they had problems with highly correlated features and/or features on different scales when computing permutation-based feature importance with random forest. Similar challenges were found by Pan [30], where the randomized input sampling explanation ignores the spatial autocorrelation of features.

In addition to the use of XAI in the case of scientific findings or to improve the understandability for end users, some challenges and future work were also identified for the optimization of machine learning models. Considering geospatial data, existing interpretation approaches [14,27,29] and machine learning models such as deep neural networks [38], CatBoost [24], or Geobost need further evaluation. Benchmark models need to be compared with newly developed models [23], e.g., XGBoost-SHAP with urban growth models such as AutoML [36]. The aim of new approaches would be to make better use of topographic and other related factors—for example, an XAI-integrated framework for physically based models and machine learning models [15], or new post-hoc XAI techniques for human-generated spatiotemporal use cases [30].

Several challenges were found in the 30 identified research studies. However, integrating XAI can lead to many more difficulties. For example, it is important to avoid overwhelming explanations, making people reluctant due to a lack of knowledge, not allowing questions, or explaining and visualizing too many variables [49]. There is also the challenge of increasing confusion and suppressing user curiosity as the complexity of XAI increases [49]. The challenges identified in research studies can assist future studies. However, care must be taken as new challenges may emerge.

#### 4. Conclusions

The following conclusion summarizes the review, provides a final view, and identifies research gaps for future work.

In this review, the research field of Explainable Artificial Intelligence for geospatial data was analyzed. In order to cover this research field by publications, Google Scholar was used in a structured way to identify relevant results from the string search. The returned Google Scholar search results were subjected to various inclusion and exclusion criteria, such as language and full-text availability. The results were then screened for relevance to this review, resulting in a total of 30 research studies, mostly from journals. The relevant research was then analyzed according to five research questions. The analysis showed that geospatial XAI is an ongoing field of research and needs further investigation.

SHAPs have been used in most of the research and have shown promising results, such as in Jena et al. [27], where they were used to select important features and understand feature contributions to model outputs. However, a map-supported presentation should also be included in the XAI part, as most of the commonly used plots cannot adequately visualize the geographical features. For example, research has shown that the summary plot for SHAP values is not a suitable visualization for the feature contribution of coordinates. A map-based presentation is essential to find and understand the impact of the geospatial component of the data. For example, successful XAI can also detect limitations and biases in training data [12]. Zimmermann et al. [25] also state that “The first usable data explainability tool is data visualization”. Even though appropriate visualization is essential, not all the identified research in this review, where the use of XAI was for the end users, considered this aspect sufficiently. It can be stated that suitable data visualization on a map needs to be included to show the geospatial context of the data. Furthermore, it can be assumed that the presentation of the findings and analysis results, as well as the explanatory information, is more important than the vast information with regard to the model and the whole background information. For example, Stadler et al. [18] provided their feature importance in a table with many features, which can make it more difficult for users to understand. A possible way to reduce complexity would be to group features instead [34].

In many cases, researchers face a trade-off between the accuracy of XAI and its explainability. For example, Veran et al. [11] and Abdollahi et al. [14] mentioned better interpretability at lower efficiency. A more complex XAI system may theoretically provide more detailed results. However, for the end user in particular, more details make the interpretation more difficult, which can lead to cognitive overload. For each use case, it has to be decided how much the explainable component can be simplified, so that the end user can still understand the system adequately, correctly, and efficiently.

This review found that geospatial XAI is currently less used for end users, rather than for model improvement. This also results in fewer evaluations of the effectiveness of the understandability of the chosen XAI techniques. Therefore, significant research efforts should be directed towards end-user XAI. Particular attention needs to be paid to geospatial visualization and to the extent to which interactive solutions could contribute to a better understanding of the machine learning results. Visualizations as well as interactive explorative methods need to be better adapted to the needs of geospatial XAI, bridging the gap between visual analytics, geospatial data, and machine learning.

Consequently, research efforts should be directed towards the evaluation of XAI techniques, including visualization approaches, to see if users understand and trust the AI system; otherwise, they will not use it [46]. Care must be taken to avoid overtrust, as this can lead to overuse, while undertrust can lead to disuse [51]. Special attention must be paid to the trade-off between the accuracy and understandability of XAI to avoid cognitive overload while still ensuring a sufficiently detailed explanation.

**Author Contributions:** Conceptualization, Cédric Roussel and Klaus Böhm; methodology, Cédric Roussel; formal analysis, Cédric Roussel; investigation, Cédric Roussel; writing—original draft preparation, Cédric Roussel; writing—review & editing, Cédric Roussel and Klaus Böhm; visualization, Cédric Roussel; supervision, Klaus Böhm; project administration, Cédric Roussel and Klaus Böhm; funding acquisition, Klaus Böhm. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Carl Zeiss Foundation, grant number P2021-02-014.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The research for this paper is part of the AI-Lab at Mainz University of Applied Sciences, which is part of the project “Trading off Non-Functional Properties of Machine Learning” at the Johannes-Gutenberg-University Mainz. The Carl Zeiss Foundation funds it.

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

## References

1. Gunning, D.; Aha, D.W. DARPA’s Explainable Artificial Intelligence Program. *AIMag* **2019**, *40*, 44–58. [CrossRef]
2. Miller, T. Explanation in artificial intelligence: Insights from the social sciences. *Artif. Intell.* **2019**, *267*, 1–38. [CrossRef]
3. Gilpin, L.H.; Bau, D.; Yuan, B.Z.; Bajwa, A.; Specter, M.; Kagal, L. Explaining Explanations: An Overview of Interpretability of Machine Learning. *arXiv* **2018**, arXiv:1806.00069. Available online: <https://arxiv.org/pdf/1806.00069> (accessed on 25 July 2023).
4. Doran, D.; Schulz, S.; Besold, T.R. What Does Explainable AI Really Mean? A New Conceptualization of Perspectives. *arXiv* **2017**, arXiv:1710.00794. Available online: <https://arxiv.org/pdf/1710.00794> (accessed on 25 July 2023).
5. Broniatowski, D.A. *Psychological Foundations of Explainability and Interpretability in Artificial Intelligence*; NIST Interagency/Internal Report (NISTIR); National Institute of Standards and Technology: Gaithersburg, MD, USA, 2021. [CrossRef]
6. Bellucci, M.; Delestre, N.; Malandain, N.; Zanni-Merk, C. Towards a terminology for a fully contextualized XAI. *Procedia Comput. Sci.* **2021**, *192*, 241–250. [CrossRef]
7. IBM. What Is Geospatial Data? Available online: <https://www.ibm.com/topics/geospatial-data> (accessed on 25 July 2023).
8. Brereton, P.; Kitchenham, B.A.; Budgen, D.; Turner, M.; Khalil, M. Lessons from applying the systematic literature review process within the software engineering domain. *J. Syst. Softw.* **2007**, *80*, 571–583. [CrossRef]
9. Jenga, K.; Catal, C.; Kar, G. Machine learning in crime prediction. *J. Ambient. Intell. Humaniz. Comput.* **2023**, *14*, 2887–2913. [CrossRef]
10. Gunning, D.; Explainable Artificial Intelligence (XAI). DARPA Information Innovation Office 2017. Available online: <https://nsarchive.gwu.edu/sites/default/files/documents/5794867/National-Security-Archive-David-Gunning-DARPA.pdf> (accessed on 5 April 2023).

11. Veran, T.; Portier, P.-E.; Fouquet, F. Crash prediction for a French highway network with an XAI-informed Bayesian hierarchical model. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 10–13 December 2020; pp. 1256–1265.
12. Matin, S.S.; Pradhan, B. Earthquake-Induced Building-Damage Mapping Using Explainable AI (XAI). *Sensors* **2021**, *21*, 4489. [CrossRef]
13. Maxwell, A.E.; Sharma, M.; Donaldson, K.A. Explainable Boosting Machines for Slope Failure Spatial Predictive Modeling. *Remote Sens.* **2021**, *13*, 4991. [CrossRef]
14. Abdollahi, A.; Pradhan, B. Urban Vegetation Mapping from Aerial Imagery Using Explainable AI (XAI). *Sensors* **2021**, *21*, 4738. [CrossRef]
15. Al-Najjar, H.A.; Pradhan, B.; Beydoun, G.; Sarkar, R.; Park, H.-J.; Alamri, A. A novel method using explainable artificial intelligence (XAI)-based Shapley Additive Explanations for spatial landslide prediction using Time-Series SAR dataset. *Gondwana Res.* **2022**, in press. [CrossRef]
16. Cilli, R.; Elia, M.; D'Este, M.; Giannico, V.; Amoroso, N.; Lombardi, A.; Pantaleo, E.; Monaco, A.; Sanesi, G.; Tangaro, S.; et al. Explainable artificial intelligence (XAI) detects wildfire occurrence in the Mediterranean countries of Southern Europe. *Sci. Rep.* **2022**, *12*, 16349. [CrossRef] [PubMed]
17. Lan, Y.; Wang, J.; Hu, W.; Kurbanov, E.; Cole, J.; Sha, J.; Jiao, Y.; Zhou, J. Spatial pattern prediction of forest wildfire susceptibility in Central Yunnan Province, China based on multivariate data. *Nat. Hazards* **2023**, *116*, 565–586. [CrossRef]
18. Stadler, S.; Betancourt, C.; Roscher, R. Explainable Machine Learning Reveals Capabilities, Redundancy, and Limitations of a Geospatial Air Quality Benchmark Dataset. *MAKE* **2022**, *4*, 150–171. [CrossRef]
19. Fang, H.; Shao, Y.; Xie, C.; Tian, B.; Shen, C.; Zhu, Y.; Guo, Y.; Yang, Y.; Chen, G.; Zhang, M. A New Approach to Spatial Landslide Susceptibility Prediction in Karst Mining Areas Based on Explainable Artificial Intelligence. *Sustainability* **2023**, *15*, 3094. [CrossRef]
20. Iban, M.C.; Bilgilioglu, S.S. Snow avalanche susceptibility mapping using novel tree-based machine learning algorithms (XGBoost, NGBoost, and LightGBM) with eXplainable Artificial Intelligence (XAI) approach. *Stoch. Environ. Res. Risk Assess.* **2023**, *37*, 2243–2270. [CrossRef]
21. Liang, L.; Daniels, J.; Phillips, R.; South, J.; Hu, L. Integrating Low-Cost Sensor Monitoring, Satellite Mapping, and Geospatial Artificial Intelligence for Intra-Urban Air Pollution Predictions. *Environ. Pollut.* **2023**, *331*, 121832. [CrossRef]
22. Mamalakis, A.; Barnes, E.A.; Ebert-Uphoff, I. Carefully Choose the Baseline: Lessons Learned from Applying XAI Attribution Methods for Regression Tasks in Geoscience. *Artif. Intell. Earth Syst.* **2023**, *2*, e220058. [CrossRef]
23. Youssef, K.; Shao, K.; Moon, S.; Bouchard, L.-S. Landslide Susceptibility Modeling by Interpretable Neural Network. *arXiv* **2023**, arXiv:2201.06837. Available online: <https://arxiv.org/pdf/2201.06837> (accessed on 6 April 2023).
24. Kim, M.; Kim, D.; Kim, G. Examining the Relationship between Land Use/Land Cover (LULC) and Land Surface Temperature (LST) Using Explainable Artificial Intelligence (XAI) Models: A Case Study of Seoul, South Korea. *Int. J. Environ. Res. Public Health* **2022**, *19*, 15926. [CrossRef]
25. Zimmermann, B.; Boussard, M.; Boulbes, N.; Grégoire, S. XAI and Geographic Information: Application to Paleoenvironmental Reconstructions. Workshop on Trustworthy Artificial Intelligence as a Part of the ECML/PKDD 22 Program, IRT SystemX, September 2022, Grenoble, France. Available online: <https://hal.science/hal-03773375/document> (accessed on 6 April 2023).
26. Bommer, P.; Kretschmer, M.; Hedström, A.; Bareeva, D.; Höhne, M.M.-C. Finding the right XAI method—A Guide for the Evaluation and Ranking of Explainable AI Methods in Climate Science. *arXiv* **2023**, arXiv:2303.00652. Available online: <https://arxiv.org/pdf/2303.00652> (accessed on 6 April 2023).
27. Jena, R.; Pradhan, B.; Gite, S.; Alamri, A.; Park, H.-J. A new method to promptly evaluate spatial earthquake probability mapping using an explainable artificial intelligence (XAI) model. *Gondwana Res.* **2022**, in press. [CrossRef]
28. Ahmed, I.; Kumara, I.; Reshadat, V.; Kayes, A.S.M.; van den Heuvel, W.-J.; Tamburri, D.A. Travel Time Prediction and Explanation with Spatio-Temporal Features: A Comparative Study. *Electronics* **2022**, *11*, 106. [CrossRef]
29. Zhang, X.; Zhao, X. Machine learning approach for spatial modeling of ridesourcing demand. *J. Transp. Geogr.* **2022**, *100*, 103310. [CrossRef]
30. Pan, M. Making Sense of Human-Generated Spatial-Temporal Data from Urban Environment. Ph.D. Thesis, The University of Iowa, Iowa City, IA, USA, 2021.
31. Li, Z. Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost. *Comput. Environ. Urban Syst.* **2022**, *96*, 101845. [CrossRef]
32. Jin, C. *Understanding Human Mobility and Urban Dynamics with Big Geospatial Data Analytics*; University of California: Santa Barbara, CA, USA, 2022.
33. Kim, M.; Kim, G. Modeling and Predicting Urban Expansion in South Korea Using Explainable Artificial Intelligence (XAI) Model. *Appl. Sci.* **2022**, *12*, 9169. [CrossRef]
34. Matuszelański, K.; Kopczewska, K. Customer Churn in Retail E-Commerce Business: Spatial and Machine Learning Approach. *JTAER* **2022**, *17*, 165–198. [CrossRef]
35. Graczyk-Kucharska, M.; Olszewski, R.; Weber, G.-W. The use of spatial data mining methods for modeling HR challenges of generation Z in greater Poland Region. *Cent. Eur. J. Oper. Res.* **2023**, *31*, 205–237. [CrossRef]



36. Kim, M.; Kim, D.; Jin, D.; Kim, G. Application of Explainable Artificial Intelligence (XAI) in Urban Growth Modeling: A Case Study of Seoul Metropolitan Area, Korea. *Land* **2023**, *12*, 420. [CrossRef]
37. Ahmed, Z.U.; Sun, K.; Shelly, M.; Mu, L. Explainable artificial intelligence (XAI) for exploring spatial variability of lung and bronchus cancer (LBC) mortality rates in the contiguous USA. *Sci. Rep.* **2021**, *11*, 24090. [CrossRef]
38. Temenos, A.; Tzortzis, I.N.; Kaselimi, M.; Rallis, I.; Doulamis, A.; Doulamis, N. Novel Insights in Spatial Epidemiology Utilizing Explainable AI (XAI) and Remote Sensing. *Remote Sens.* **2022**, *14*, 3074. [CrossRef]
39. Pradhan, B.; Jena, R.; Talukdar, D.; Mohanty, M.; Sahu, B.K.; Raul, A.K.; Abdul Maulud, K.N. A New Method to Evaluate Gold Mineralisation-Potential Mapping Using Deep Learning and an Explainable Artificial Intelligence (XAI) Model. *Remote Sens.* **2022**, *14*, 4486. [CrossRef]
40. Sachit, M.S.; Shafri, H.Z.M.; Abdullah, A.F.; Rafie, A.S.M.; Gibril, M.B.A. Global Spatial Suitability Mapping of Wind and Solar Systems Using an Explainable AI-Based Approach. *IJGI* **2022**, *11*, 422. [CrossRef]
41. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
42. Chen, T.; Guestrin, C. XGBoost. *arXiv* **2016**, arXiv:1603.02754. Available online: <https://arxiv.org/pdf/1603.02754> (accessed on 20 April 2023).
43. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 3146.
44. Zhang, Z. Boosting Algorithms Explained. *Towards Data Science [Online]*, June 26, 2019. Available online: <https://towardsdatascience.com/boosting-algorithms-explained-d38f56ef3f30> (accessed on 20 April 2023).
45. Lundberg, S.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. *arXiv* **2017**, arXiv:1705.07874. Available online: <https://arxiv.org/pdf/1705.07874> (accessed on 24 April 2023).
46. Ribeiro, M.T.; Singh, S.; Guestrin, C. “Why Should I Trust You?”: Explaining the Predictions of Any Classifier. *arXiv* **2016**, arXiv:1602.04938. Available online: <https://arxiv.org/pdf/1602.04938> (accessed on 2 May 2023).
47. Doshi-Velez, F.; Kim, B. Towards a Rigorous Science of Interpretable Machine Learning. *arXiv* **2017**, arXiv:1702.08608. Available online: <https://arxiv.org/pdf/1702.08608> (accessed on 11 April 2023).
48. Carvalho, D.V.; Pereira, E.M.; Cardoso, J.S. Machine Learning Interpretability: A Survey on Methods and Metrics. *Electronics* **2019**, *8*, 832. [CrossRef]
49. Tosun, A.B.; Pullara, F.; Becich, M.J.; Taylor, D.L.; Chennubhotla, S.C.; Fine, J.L. HistoMapr™: An Explainable AI (xAI) Platform for Computational Pathology Solutions. In *Artificial Intelligence and Machine Learning for Digital Pathology*; Springer: Cham, Switzerland, 2020; pp. 204–227.
50. Xing, J.; Sieber, R. The challenges of integrating explainable artificial intelligence into GeoAI. *Trans. GIS* **2023**, *27*, 626–645. [CrossRef]
51. Parasuraman, R.; Riley, V. Humans and Automation: Use, Misuse, Disuse, Abuse. *Hum. Factors* **1997**, *39*, 230–253. [CrossRef]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.