[AM-08-108] Explainable AI Applications in GIS

Abstract

With the increasing availability of training data and the advancement of computer hardware technologies, such as graphic processing units (GPUs) and storage devices, artificial intelligence (AI) techniques have been widely applied in multiple fields due to their excellent performance and transferability. In addition to the benefits AI has brought to the industry, it has had a huge impact on conventional methods in various fields of academia. It has also given rise to some new cross-disciplinary fields, such as geospatial artificial intelligence (GeoAI). However, since some AI methods represented by deep learning have massive parameters and complex network structures, their model decision-making basis is often difficult to understand (i.e., black box), which affects the AI method promotion in high-risk fields (e.g., autonomous driving). To make the “black box” models more transparent, AI scientists have proposed some methods to explain AI models, i.e., explainable artificial intelligence (XAI) methods. This entry briefly presents the main XAI methods, provides three examples of applying XAI methods to geographical information science (GIS) research, and summarizes the direction of related studies.

Author & citation

Cheng, X. (2025). Explainable AI Applications in GIS. The Geographic Information Science & Technology Body of Knowledge (Issue 2, 2025 Edition), John P. Wilson (ed.).
DOI: 10.22224/gistbok/2025.2.1.

Explanation

1. Explainable Artificial Intelligence (XAI)
2. XAI Applications in GIS
3. Summary and Direction

1. Explainable Artificial Intelligence (XAI)

With the development and widespread applications of artificial intelligence (AI) techniques in recent years, explainable artificial intelligence (XAI) targeting the “black box” characteristics of AI has also received more attention from scientists and users. However, XAI is not a new term proposed recently. Its history can be traced back to the studies of expert systems in the 1980s (Xu et al., 2019). At that time, the research was trying to convert the knowledge from human experts into rules that the system could understand and operate according to, so that the working principle of the expert system is human understandable (Ali et al., 2023).

Since AI technologies have strong transferability and can perform better than traditional methods in multiple tasks, AI has been widely applied in the GIS field. GIS scientists have transferred and developed new AI methods based on specific characteristics (e.g., spatio-temporal dependence) and application requirements to solve research problems in the field, thereby promoting the emergence and development of interdisciplinary–geospatial artificial intelligence (GeoAI) (Janowicz et al., 2020). However, GeoAI also has the “black

box" characteristics inherited from general AI, so GIS scientists have made attempts to open the black boxes. This entry aims to provide basic information on XAI to researchers and students in the GIS field and to promote the application of XAI techniques in GIS. Figure 1 displays the scope of this entry.

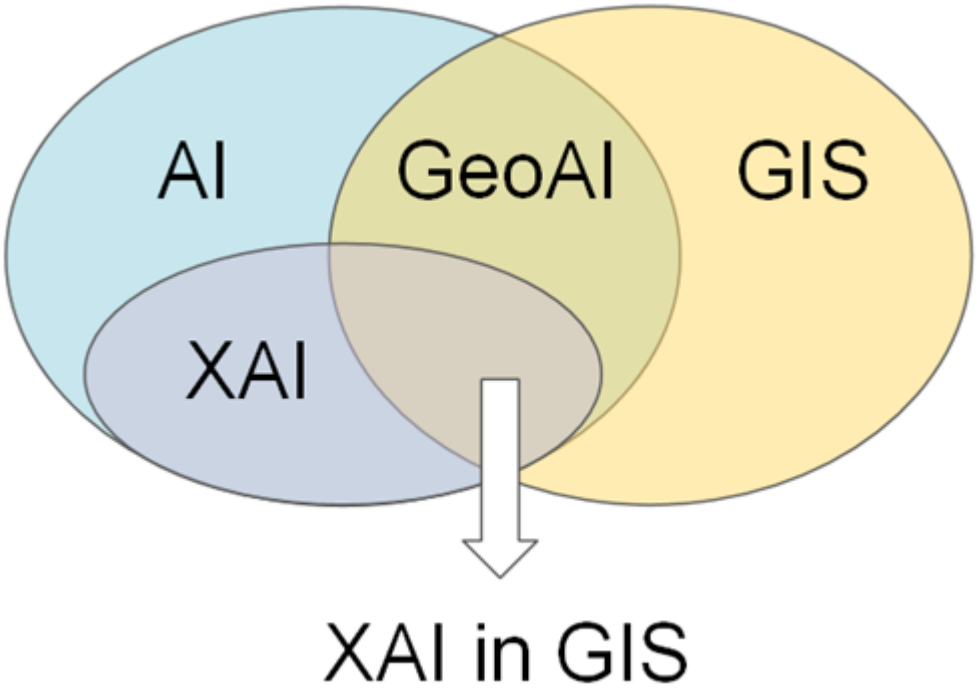


Figure 1. The scope of XAI applications in GIS for this Body of Knowledge entry. Source: author.

1.1 Main XAI Approaches

The scope of XAI approaches in a broad sense contains many methods, such as interpretable machine learning methods (e.g., decision trees, K-means, and linear regression), and visualization-based methods that optimize the display of model parameters, structures, and intermediate results. Due to limited length, this entry focuses on a narrower but more acceptable scope of XAI approaches, i.e., the XAI methods for explaining deep learning techniques (e.g., deep neural networks). XAI methods have a variety of taxonomies from different perspectives (Ali et al., 2023). Some XAI methods can only be applied to specific models (e.g., neural networks), while others can be applied to any model because they just need to observe the corresponding changes in the input and output of the model. The former is model-specific [e.g., Gradient-weighted Class Activation Mapping (Grad-CAM) and Layer-wise Relevance Propagation (LRP)] and the latter is model-agnostic [e.g., Local Interpretable Model Agnostic explanations (LIME) and SHapley Additive exPlanations (SHAP)]. There is also a taxonomy that classifies XAI methods according to their principles (e.g., backpropagation-based and perturbation-based). Among all taxonomies, the most widely accepted taxonomy of XAI methods divides the methods according to the scope of the obtained explanation, i.e., local explanation and global explanation. The former refers to XAI methods that can explain a single sample or prediction, while the latter refers to XAI methods that can explain the global task or the

entire dataset.

XAI has been developed for many years and is still a hot field. This entry cannot list all proposed XAI methods and their improved versions. Due to the limitations of the main text and the number of references, this entry takes an example to list the names of the main XAI methods. Refer to Cheng et al. (2023), global explanation XAI methods include Activation Maximization and Conceptual Sensitivity; local explanation XAI methods include LIME, Occlusion, Perturbation, SHAP, Saliency, Integrated Gradients, Grad-CAM, SmoothGrad, LRP, and Deep Taylor Decomposition (DTD). For more detailed information on XAI methods, the reader could refer to Ali et al. (2023) and Samek et al. (2021), as well as the original articles of each XAI method.

1.2 Benefits of Applying XAI Methods

Compared with conventional models, AI techniques are more capable of solving complex nonlinear problems. However, AI represented by deep learning has huge parameters and complex connection structures between neurons in multiple layers. Hence, the basis of model decision-making is difficult for people to understand. It creates the dilemma of choosing between model performance and transparency when applying AI methods. The XAI methods can break this dilemma somewhat. The main benefits of applying XAI methods are summarized below. First, XAI can provide users with explanations of the basis for AI model decisions, just as opening a transparent window on a black box, allowing users to peek into the model's operating mechanism through the explanations obtained by applying XAI (Dramsch et al., 2025). If the decision basis of AI models is consistent with the users' expectations or domain prior knowledge, the users will trust the model and can apply the trained AI model to actual scenarios. In other cases, the Clever Hans phenomenon may occur, i.e., the AI model can achieve the specific task well based on the unexpected decision-making basis (Lapuschkin et al., 2019). XAI can also help users to discover such phenomena based on the explanations of AI models. Figure 2 displays an example where the Clever Hans phenomenon can be found by XAI explanations.

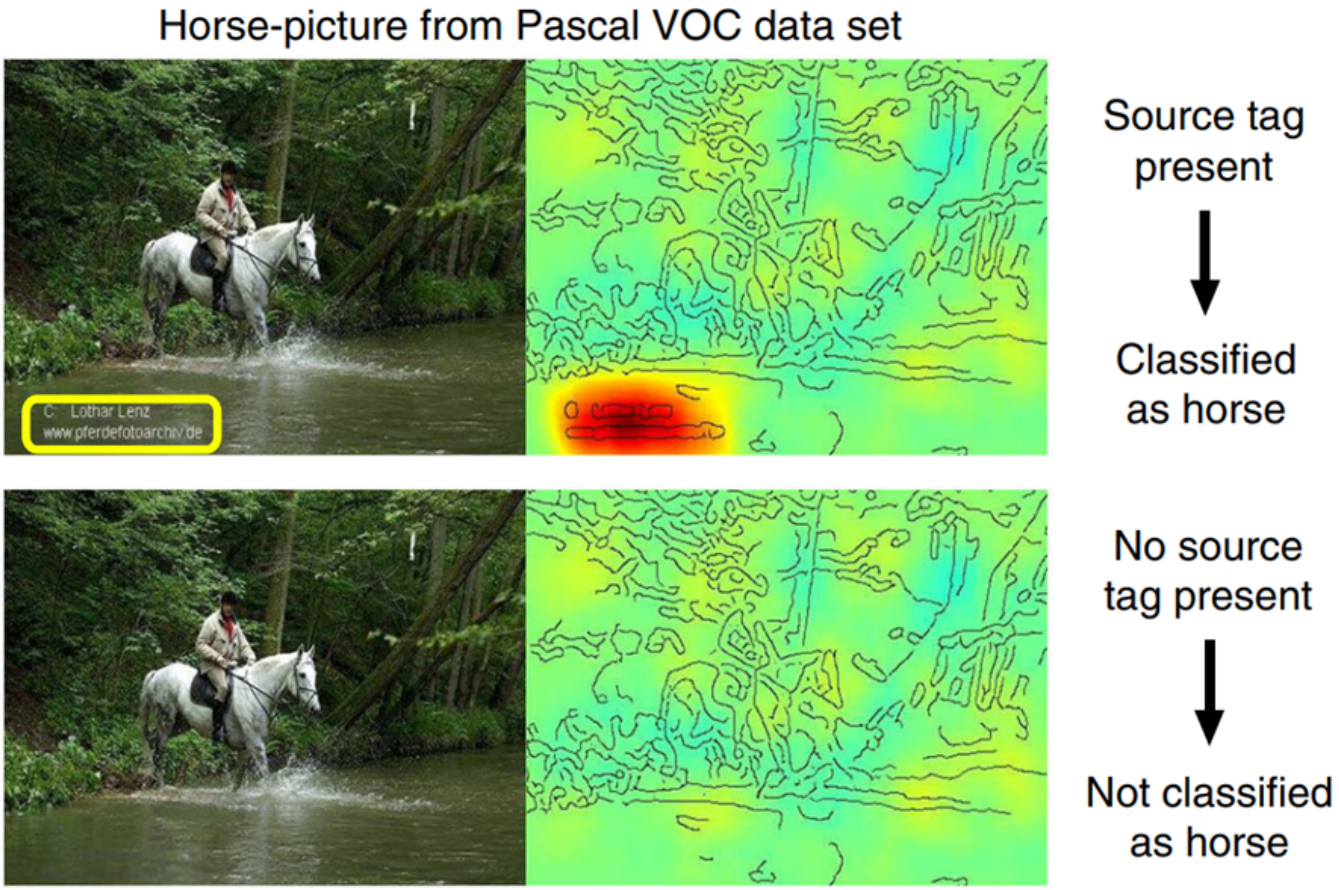


Figure 2. An example of the Clever Hans phenomenon. An image is identified as a “horse” by an AI model, but the decision basis of the model is the source tag information presented in the image. The right images show the XAI explanations. The pixels in red color indicate that they are important to the decision of the AI model. Source: Lapuschkin et al. (2019).

Second, XAI allows scientists to make further model improvements. By reorganizing training data or intervening in model training, the decision-making basis of AI models can be controlled to be consistent or close to human expectations. In this way, AI techniques can perform excellently on specific target tasks while ensuring that the decision basis is human-understandable. For instance, Cheng et al. (2022) has shown that the explanations of models by XAI can guide users to intervene in model training by expertise, thus improving the AI model performance on multiple tasks based on leaf image data. Anders et al. (2022) provides more information on improving AI models using XAI methods. Last, XAI has the potential to discover new knowledge on specific tasks. That is, scientists can be inspired to conduct further experiments for study based on the explanations of AI models by XAI (e.g., which input variables have a great impact on the target variable), thereby promoting the development of the discipline. Unlike conventional models designed and operated based on known expertise, AI can automatically discover and learn the complex relationship between input and output data to achieve specific tasks. That mined knowledge can be new to humans. Moreover, the excellent performance of AI models has shown that computers have surpassed humans in multiple tasks (e.g., image classification). It means that in specific tasks, especially in emerging fields that have not been fully investigated (e.g., Go game), humans can even learn from computers based on the explanations of AI models.

2. XAI Applications in GIS

To make the AI or GeoAI models transparent, one typical way is to develop spatially-explicit GeoAI models (Liu & Biljecki, 2022), which integrate domain expertise into model training or training data organization. For instance, Liu et al. (2022) designed the connection structure of the neural networks based on the upstream and downstream information of the river network. Another way is to use the XAI methods to explain AI or GeoAI models applied in GIS studies. This entry lists three examples with different research purposes for a brief introduction.

2.1. Case #1: XAI application driven by task-specific knowledge mining in spatio-temporal dimensions.

AI techniques can automatically mine high-dimensional data without the guidance of experts. Like deep learning models, they discover and learn knowledge during training based on input data and corresponding output data, and then perform tasks. XAI can explain AI model decisions and transfer the discovered knowledge to a human-understandable form. Cheng et al. (2021) proposed a method–Spatio-temporal Layer-wise Relevance Propagation (ST-LRP), which aims to discover task-specific knowledge by actively designing classification tasks, training deep learning models to mine spatio-temporal data, and then obtaining and displaying the results using XAI (i.e., the LRP algorithm), as shown in Figure 3.

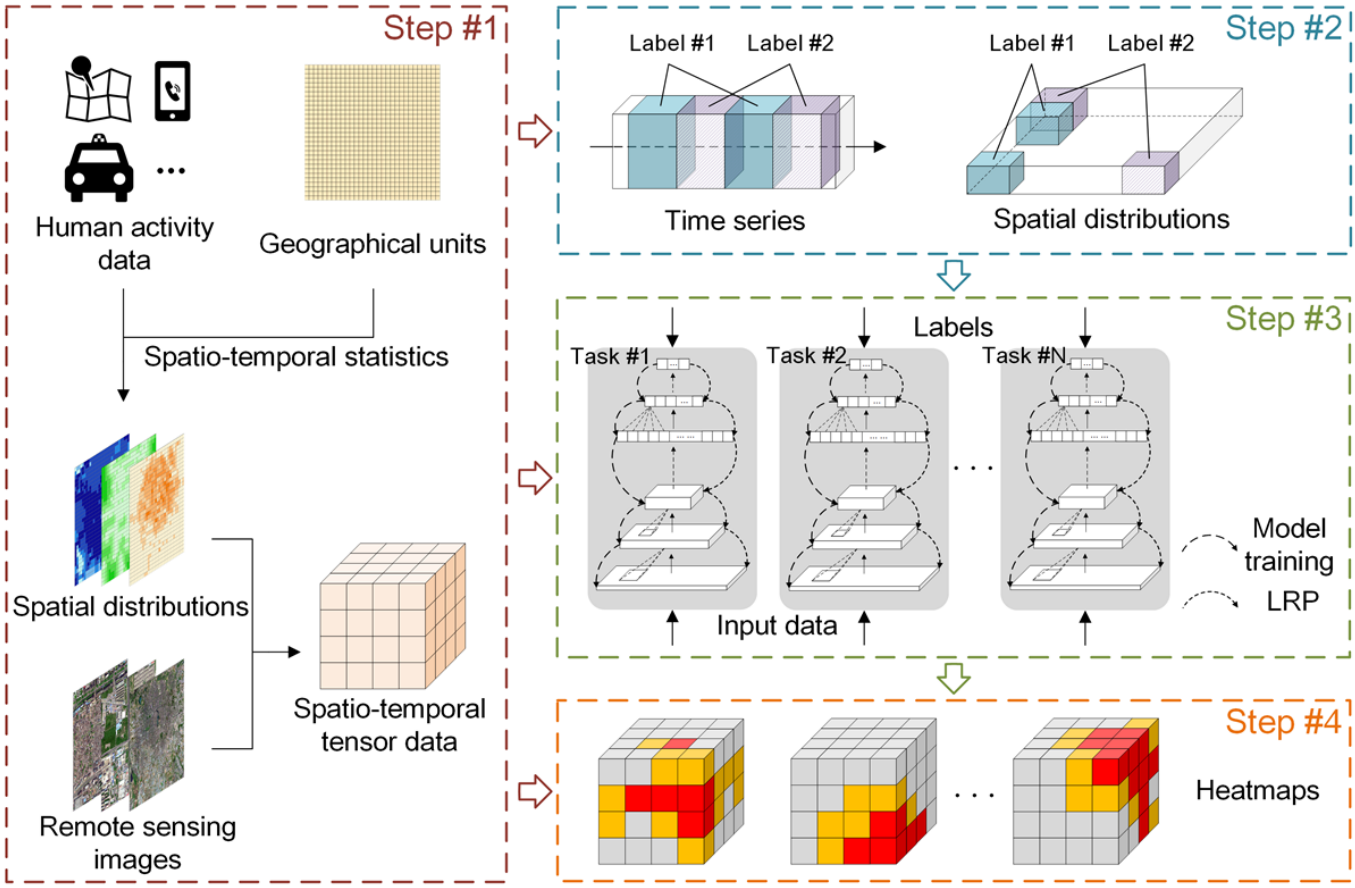


Figure 3. The procedure of the ST-LRP method with four steps: data organization, data labeling, model training and XAI application, and explanation visualization and analysis.

Source: Cheng et al. (2021).

The ST-LRP method applies to multiple spatio-temporal datasets. The raw data are transformed into tensor form according to spatio-temporal information. Due to the difference between GIS and AI fields, local explanations of individual samples in GeoAI contain spatio-temporal information, which makes it difficult to summarize them into global explanations and is not conducive to knowledge discovery. Therefore, before implementing the XAI method, it is necessary to ensure that one of the temporal and spatial dimensions of labeled tensor data is fixed or can be summarized (e.g., with the same study area). In Cheng et al. (2021), the Beijing taxi trajectories are processed into spatio-temporal distributions of origin and destination points (OD) as the experimental data. A deep learning model is trained to discover knowledge about taxi OD changes on weekdays and weekends/holidays. Figure 4 displays the model explanations in temporal and spatial dimensions by XAI. The original paper (Cheng et al., 2021) provides the verification and detailed analysis of the explanation results. This study shows that XAI can benefit knowledge discovery in GIS.

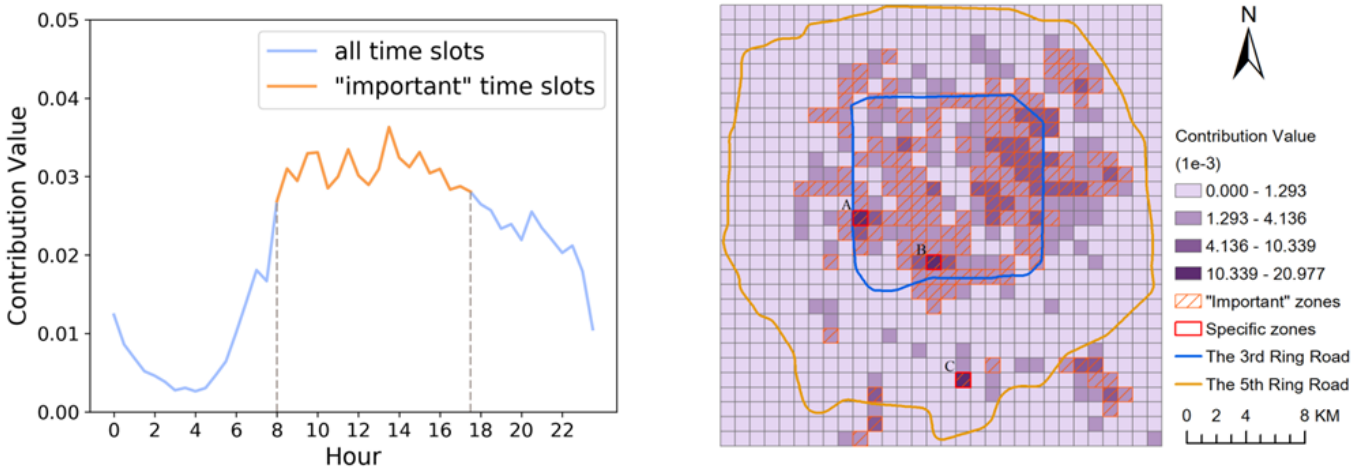


Figure 4. An example of knowledge discovery results by ST-LRP (left: results in temporal dimension; right: results in spatial dimension). The results reflect which spatio-temporal units are important to distinguish Beijing taxi ODs between weekdays and weekends/holidays. Source: Cheng et al. (2021).

2.2. Case #2: XAI explanation of spatial effect extraction using machine learning methods.

Since GIS scientists have applied AI techniques to GIS or GeoAI studies, there has been competition between conventional and AI methods. The model design and operation of traditional methods have a sufficient mathematical basis. Although AI methods can outperform traditional GIS methods in performance on many tasks, they are usually regarded as black boxes, and their decision-making basis cannot be directly observed and understood by users, which hinders their application and development in GIS. Li (2022) took the spatial lag, one of the typical spatial effects, as an example for studying. According to formulas with known forms and parameters (i.e., Spatial Lag Model), Li (2022) generated input and output variables with known spatial lag, and then used a popular machine

learning method–XGBoost to train based on the simulation data. After model training, a model-agnostic XAI method–SHAP was used to explain the trained model and visualize the knowledge learned by XGBoost.

Figure 5 shows an example of a comparison between the XAI explanations with the true values of the simulation data. The two groups of results (i.e., formula parameters and spatial lag) are close, which proves that AI methods can learn the knowledge on spatial effects (i.e., spatial lag), thereby increasing the confidence of AI models and promoting their development in GIS.

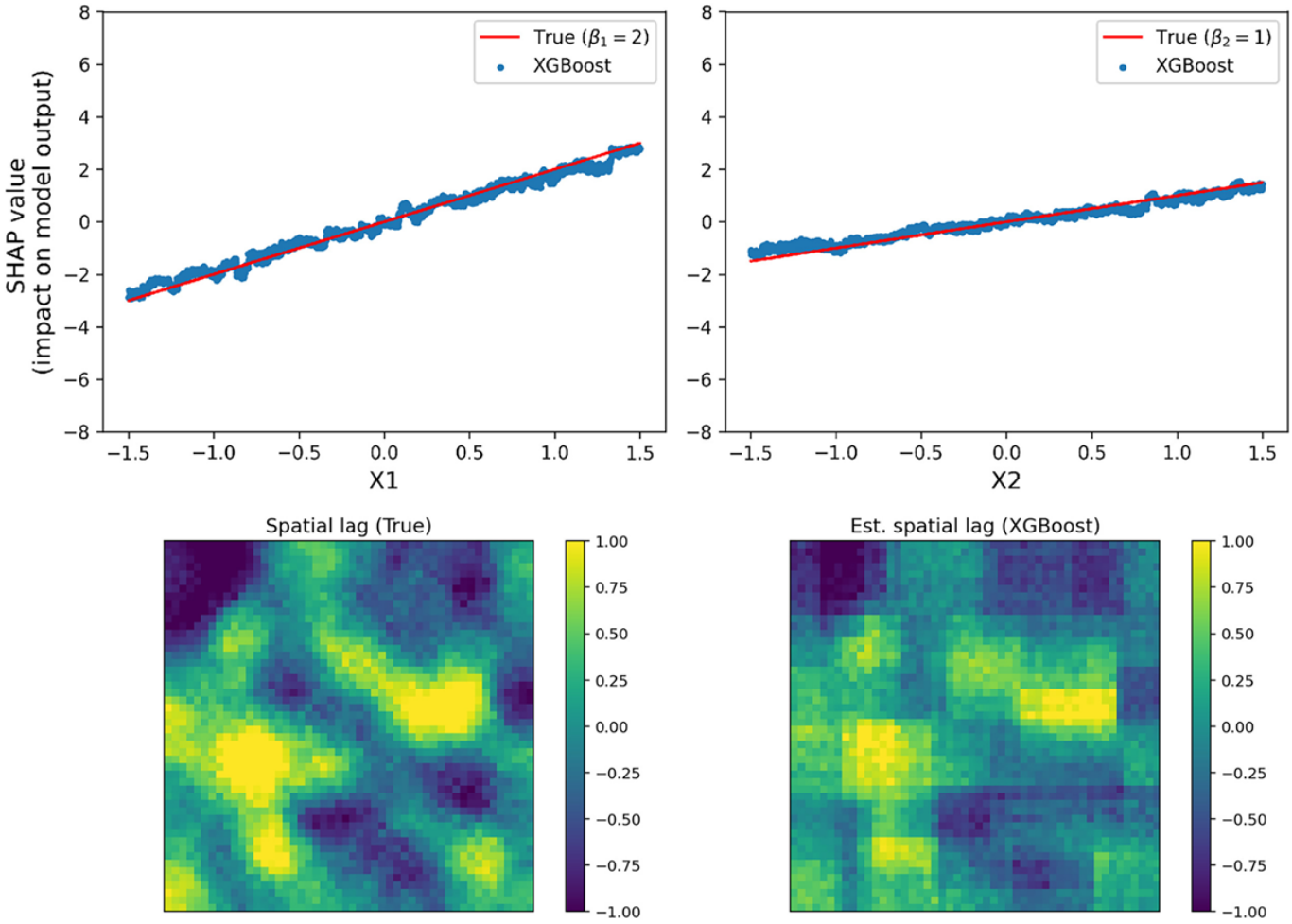


Figure 5. Comparison between true values (i.e., known formula parameters and spatial lag) with explanations of machine learning methods by SHAP. Source: Li (2022).

2.3. Case #3: Comparison of explanations using three XAI methods in an object image classification task.

It is difficult to make a comprehensive comparison between model explanations using multiple XAI methods, even in general AI research. Multiple XAI methods have different principles and explain the decision-making basis of the AI model from various aspects. Some explanations are perhaps consistent with experience and some are unexpected to users. Since the decision-making basis of AI models has no “true value”, it is unfair for humans to subjectively define whether the explanation is right or wrong. The evaluation of model explanation by XAI in GIS studies is more difficult than in general AI (e.g., to obtain

global explanations). Hsu & Li (2023) innovatively used an image dataset originally for object detection to train a deep learning model for a terrain classification task. In that case, the location and range of the terrain features in the image are known, which provides a direct and clear benchmark for the evaluation of explanations by XAI. The data used by Hsu & Li (2023) include eight types of terrain features. After training a deep learning model for the terrain classification task, three XAI methods (i.e., Occlusion, Grad-CAM, and Integrated Gradients) were used to explain the trained model. Figure 6 shows an example of the local explanation comparison of the “volcano” terrain object. The explanation results using three XAI methods are different, and the information of “true” terrain objects (i.e., blue rectangles) provides a benchmark for comparison. Hsu & Li (2023) compared the model explanations by XAI methods in GIS studies. It highlights that multiple XAI methods may have different explanation results for the same trained model and targeted task. In addition, the clever use of an object detection dataset for an image classification task also guides the related GIS studies based on image data.

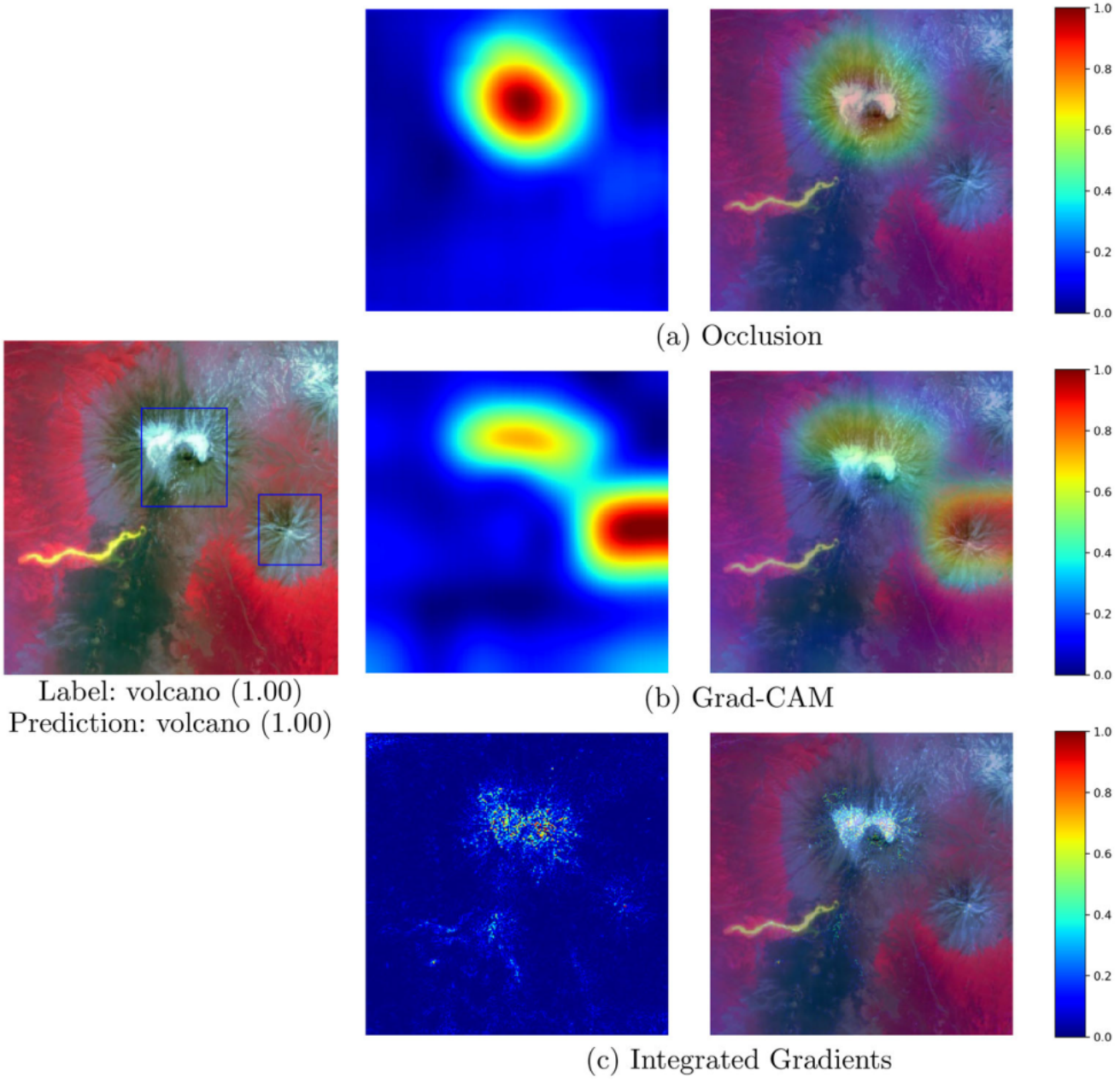


Figure 6. An example of model explanations using three XAI methods. The larger the pixel value in the results, the more important it is for image classification. The blue rectangles in

the original images display the “true value” of the model explanation. Source: Hsu & Li (2023).

3. Summary and Discussion

With the development of AI and GeoAI methods, the XAI application in GIS will attract more attention from scientists. However, the problems that current XAI methods have in general AI applications also exist in GIS. These include lacking widely accepted datasets or tools for evaluating XAI methods [e.g., Quantus (Hedstrom et al., 2023)], XAI explanations (e.g., pixel-based attribution maps) cannot be understood by non-AI expert users, and the requirements for applicability, efficiency, and robustness of the XAI methods themselves. In addition, the GIS field has extra requirements for the XAI application (Cheng et al., 2023), such as diverse data types (e.g., taxi trajectories and street view images), non-independent variables due to spatio-temporal dependencies, and data unavailability due to missing or privacy protection. This entry aims to introduce basic information about XAI to people who majored in GIS and promote the application and improvement of XAI methods.

References

Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Dıaz-Rodriguez, N., & Herrera, F. (2023). Explainable artificial intelligence (XAI): What we know and what is left to attain Trustworthy Artificial Intelligence. Information Fusion, 99 , 101805.

Anders, C. J., Weber, L., Neumann, D., Samek, W., Muller, K.-R., & Lapuschkin, S. (2022). Finding and removing Clever Hans: Using explanation methods to debug and improve deep models. Information Fusion, 77 , 261– 295.

Cheng, X., Doosthosseini, A., & Kunkel, J. (2022). Improve the deep learning models in forestry based on explanations and expertise. Frontiers in Plant Science, 13 , 902105.

Cheng, X., Vischer, M., Schellin, Z., Arras, L., Kuglitsch, M. M., Samek, W., & Ma, J. (2023). Explainability in GeoAI. In Handbook of Geospatial Artificial Intelligence (pp. 177–200). CRC Press.

Cheng, X., Wang, J., Li, H., Zhang, Y., Wu, L., & Liu, Y. (2021). A method to evaluate task-specific importance of spatio-temporal units based on explainable artificial intelligence. International Journal of Geographical Information Science, 35 , 2002–2025.

Dramsch, J. S., Kuglitsch, M. M., Fernandez-Torres, M.-A., Toreti, A., Albayrak, R. A., Nava, L., Ghaffarian, S., Cheng, X., Ma, J., Samek, W. et al. (2025). Explainability can foster trust in artificial intelligence in geoscience. Nature Geoscience, (pp. 1–3).

Hedstrom, A., Weber, L., Krakowczyk, D., Bareeva, D., Motzkus, F., Samek, W., Lapuschkin, S., & Ho¨hne, M. M.-C. (2023). Quantus: An explainable AI toolkit for responsible evaluation of neural network explanations and beyond. Journal of Machine Learning

Research, 24 , 1–11.

Hsu, C.-Y., & Li, W. (2023). Explainable GeoAI: Can saliency maps help interpret artificial intelligence's learning process? An empirical study on natural feature detection. International Journal of Geographical Information Science, 37 , 963–987.

Janowicz, K., Gao, S., McKenzie, G., Hu, Y., & Bhaduri, B. (2020). GeoAI: Spatially explicit artificial intelligence techniques for geographic knowledge discovery and beyond. International Journal of Geographical Information Science, 0(0), 1-13.

Lapuschkin, S., Waldchen, S., Binder, A., Montavon, G., Samek, W., & Muller, K.-R. (2019). Unmasking Clever Hans predictors and assessing what machines really learn. Nature Communications, 10 , 1096.

Li, Z. (2022). Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost. Computers, Environment and Urban Systems, 96 , 101845.

Liu, P., & Biljecki, F. (2022). A review of spatially-explicit GeoAI applications in Urban Geography. International Journal of Applied Earth Observation and Geoinformation, 112 , 102936.

Liu, Y., Hou, G., Huang, F., Qin, H., Wang, B., & Yi, L. (2022). Directed graph deep neural network for multi-step daily streamflow forecasting. Journal of Hydrology, 607 , 127515.

Samek, W., Montavon, G., Lapuschkin, S., Anders, C. J., & Muller, K.-R. (2021). Explaining deep neural networks and beyond: A review of methods and applications. Proceedings of the IEEE, 109 , 247–278.

Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D., & Zhu, J. (2019). Explainable AI: A brief survey on history, research areas, approaches and challenges. In Natural Language Processing and Chinese Computing: 8th CCF International Conference, NLPCC 2019, Dunhuang, China, October 9–14, 2019, proceedings, part II 8 (pp. 563–574). Springer.