ARTIFICIAL INTELLIGENCE IN GEOGRAPHICAL RESEARCH: FROM SPATIAL OBSERVATION TO INFORMED ACTION - A SYNTHESIS OF THE AIAG STUDIES AND THE GEOGRAPHICAL FOUNDATIONS OF GEOAI

Zsolt MAGYARI-SÁSKA[1]

An Afterword from the Co-editor

DOI: 10.21163/GT_2026.212.13

ABSTRACT

This paper synthesises the twelve contributions included in the AIAG special issue and considers their findings within the broader development of geospatial artificial intelligence. The studies show how AI and related computational methods can support the identification of geographical phenomena, the estimation of environmental conditions in data-sparse areas, the analysis of spatiotemporal processes, the construction of land-use and hazard scenarios, and the integration of data-driven models with physical knowledge. Rather than treating algorithms as the principal outcome, the synthesis focuses on the geographical evidence they produce, including flood boundaries, pollution estimates, rainfall forecasts, biomass inventories, burned-area maps, land-use projections and susceptibility assessments. The discussion then shifts from what AI performs in the individual studies to what geography contributes to AI. It examines how geographical research defines meaningful entities, represents scale, connectivity and temporal structure, organises heterogeneous evidence, keeps environmental and social processes visible, and connects model outputs to places, institutions and decisions. These contributions are essential because spatial coordinates alone do not make a model geographically meaningful.

The paper also considers the conditions required for reliable GeoAI, particularly spatially appropriate validation, transferability between regions, interpretability and the spatial communication of uncertainty. It argues that AI is most useful when it complements GIS, remote sensing, field observations, statistical analysis, physical modelling and expert knowledge. Its contribution lies not in replacing geographical judgement, but in extending the capacity to observe, estimate and anticipate spatial processes while keeping model outputs connected to the geographical contexts and decisions they represent.

Keywords: *GeoAI; geographical knowledge; spatial dependence; geographical modelling; decision support; spatial validation; explainable AI*

1. INTRODUCTION

Geospatial artificial intelligence (GeoAI) represents an evolving set of analytical methodologies that extend the capacity of geographers to observe, measure, model, and anticipate spatial phenomena. The twelve contributions assembled in this special issue demonstrate this analytical progression across diverse environmental and socio-economic contexts:

[1] *Babeș-Bolyai University, Faculty of Geography, 400006 Cluj-Napoca, Romania, zsolt.magyari@ubbcluj.ro co-editor of AIAG.*

1. Kritchayan INTARAT, Phatsachon KUANKHAMNUAN and Woraman JANGSAWANG / *Flood Mapping in Phra Nakhon Si Ayutthaya, Thailand, Utilizing Sentinel-1 SAR Imagery and Deep Learning Approaches*
2. Triyatno TRIYATNO, Lailatur RAHMI, Syafri ANWAR, Beni AULIA and Sumayyah Aimi Mohd NAJIB / *Projecting LULC Change (2025–2039) and Modeling Landslide Susceptibility Using Cellular Automata–Markov Chain (CA–MC) in the Kambang Watershed, West Sumatra, Indonesia*
3. Sri NURDIATI, Mochamad Tito JULIANTO, Ionel HAIDU, Muhammad Daryl FAUZAN, Hari NURDIANTO, Syukri Arif RAFHIDA and Mohamad Khoirun NAJIB / *Predicting Fire Hotspots in Kalimantan Based on Climate Factors Using Gaussian Process Regression and Long Short-Term Memory*
4. Uca SIDENG, Nurul Afdal HARIS and Mustari S. LAMADA / *An Integrated ANN–CA Model for Land-Use Change Prediction and Flood-Risk Mitigation: A Case Study of Enrekang Regency, Indonesia*
5. Ananto A. AJI, Syaiful Muflichin PURNAMA and Vina Nurul HUSNA / *Random Forest–Based Mapping of Riverine Plastic Pollution from Sentinel-2 in Google Earth Engine: LULC and Rainfall Controls in Kendal Regency, Indonesia*
6. Jumadi JUMADI, Kuswaji Dwi PRIYONO, Ali Hasan ABDULLAH, Supari SUPARI, Hamza AIT ZAMZAMI, Farha SATTAR, Muhammad NAWAZ, Hamzah HASYIM and Steve CARVER / *Performance of Machine-Learning and Deep-Learning Models for Predicting Rainfall in a Large Watershed: Case Study of the Bengawan Solo River Basin, Indonesia*
7. Imad KATIBA, Hanae BELMAJDOUB and Khalid MINAOUI / *Hybrid Deep-Learning and Reinforcement Framework for Physically Consistent Precipitation Control along the Moroccan–Iberian Coasts*
8. Maliwan THABTHONG, Teerawong LAOSUWAN, Satith SANGPRADID, Yannawut UTTARUK, Jenjira PIAMDEE, Piyatida AWICHIN, Titipong PHOOPHATHONG and Maharaja SINGHARAJ / *An AI-Based Framework for Estimating $PM_{2.5}$ Using Satellite Aerosol Optical Depth and Meteorological Data in a Coastal Industrial Region of Thailand*
9. Pimpilai WUNAPHAI, Teerawong LAOSUWAN, Satith SANGPRADID, Yannawut UTTARUK, Narueset PRASERTSRI, Thinnakon ANGKAHAD and Piyatida AWICHIN / *Geospatial Deep Learning for Biomass and Carbon-Stock Estimation Using UAV-Derived Canopy Metrics*
10. Worawit SUPPAWIMUT and Ratchaphon SAMPHUTTHANONT / *Deep Learning–Based Detection of Agricultural Burned Areas Using True-Color Satellite Imagery*
11. Thanh Long NGUYEN, Thi Ngoc Ha TRAN, Quoc Lap KIEU and Dieu Trinh NGUYEN / *A Shear-Strength-Derived Lithological Weakness Index for GIS-Based Landslide-Susceptibility Modelling in Mountainous Road Corridors of Northern Vietnam*
12. Sorawit ONPHUA, Sakhan TEEJUNTUK and Chakrit NA TAKUATHUNG / *Multi-Algorithm Classification and CA–Markov Prediction of Land-Use Change across Three Temporal Intervals: A Case Study of the Loe Tor Royal Project, Tak Province, Thailand*

These studies address diverse environmental phenomena, including riverine flooding, precipitation patterns, forest fires, atmospheric pollution, riverine plastic pollution, plantation biomass, land-use transitions, and slope instability. Study areas span alluvial floodplains, tropical river basins, industrial coastal zones, commercial plantations, and mountainous transport corridors. Collectively, these investigations examine complex spatial and temporal systems by integrating multi-source geospatial data, computational models, and domain knowledge of localized physical processes.

The papers employ a broad spectrum of AI and related computational techniques, including semantic image segmentation, supervised classification, nonlinear regression, recurrent neural networks, ensemble tree architectures, cellular automata, Markov chains, and reinforcement learning. The algorithmic development is rarely the primary objective; the principal outputs consist of empirical spatial evidence, such as mapped flood extents, aerosol concentration fields, rainfall forecasts, burned scar geometries, land-use trajectories, biomass inventories, and susceptibility zonations. The geographical value of AI lies in this conversion of heterogeneous observations into spatially explicit evidence.

In the following sections, numbers in square brackets refer to the twelve AIAG contributions listed above.

2. THEMATIC SYNTHESIS OF THE AIAG STUDIES

2.1. Seeing geographical phenomena more clearly

A first group of studies uses AI to map features and processes that are difficult to identify consistently through conventional interpretation. This is especially relevant to satellite and UAV imagery, where the growing volume of observations has made manual analysis increasingly impractical. Remote-sensing archives provide repeated coverage of large areas, but the images still need to be classified, segmented and interpreted in relation to conditions on the ground. Intarat, Kuankhamnuan and Jangsawang [1] examine this problem through flood mapping in Phra Nakhon Si Ayutthaya Province, Thailand. They combine Sentinel-1 synthetic aperture radar imagery with a U-Net segmentation model using different ResNet encoders. The model distinguishes flooded from non-flooded pixels and produces continuous maps of the inundated area. Sentinel-1 data are well suited to this task because radar observations remain available under cloud cover and at night, when optical imagery may be unavailable or difficult to interpret. The comparison of the three encoders also shows that increasing network depth does not necessarily improve the result. ResNet101 provided a better balance between segmentation accuracy, generalisation and computational demand than ResNet50 or ResNet152. Maps produced soon after a flood can support the identification of affected settlements, roads and agricultural land and can be compared with the spatial extent of earlier events. Their wider operational use, however, depends on processing speed, computing capacity and validation across several floods and geographical settings. Good performance within the original study area does not guarantee that the same model will work equally well elsewhere, since part of what it learns is specific to the landscape and the radar characteristics of the training data.

Suppawimut and Samphutthanont [10] investigate the detection of agricultural burning in rice-growing areas of northern Thailand. Reference burned-area masks are first derived from the differenced Normalized Burn Ratio, after which a U-Net model is trained to identify burned patches from true-colour satellite imagery. The spectral index is therefore used to prepare the training data, while the resulting model operates with simpler RGB inputs. A particularly relevant part of the study is the spatial separation of the training, validation and test areas. Nearby pixels and image tiles often resemble one another, so random sampling can allow almost identical spatial patterns to appear in both the training and test datasets, leading to overly optimistic estimates of model performance. The spatial hold-out design provides a more realistic assessment of how well the model performs in previously unseen areas. The resulting burn maps can also show where agricultural fires are concentrated and where they may contribute to seasonal haze and fine-particle pollution. In this way, the method provides spatial information that can support the targeting of mitigation measures without treating image classification as an end in itself.

Wunaphai and colleagues [9] work at a finer spatial scale, combining UAV imagery, canopy-height models, field measurements and deep-learning-based tree detection to estimate biomass and carbon stocks in rubber plantations. Tree height, crown dimensions and stand density are derived from the UAV data, while the field plots provide the biomass measurements needed for calibration. Deep learning is mainly used to identify individual trees and organise the canopy information into a plantation inventory. The value of the approach lies in the connection it creates between a relatively small number of field plots and a much more detailed spatial representation of the plantation. Field observations therefore remain essential; the model does not replace them, but extends their usefulness beyond the locations where measurements were collected.

2.2. Estimating what cannot be measured everywhere

A second group of studies uses AI to estimate environmental conditions in areas where direct observations are limited. This is a common problem in geographical research: monitoring stations are unevenly distributed, field measurements are costly, and conditions may vary considerably within the same administrative region. Models can partly address these gaps by combining observations from several sources and estimating values for locations where no direct measurements are available.

Thabthong and colleagues [8] apply this approach to the estimation of daily $PM_{2.5}$ concentrations in Rayong Province, a coastal and industrial area of Thailand. They combine measurements from ground stations with satellite-derived aerosol optical depth and meteorological variables obtained from reanalysis data, including temperature, relative humidity, wind speed and boundary-layer height. Random Forest performed best among the methods tested. The analysis also identified wind speed, humidity and boundary-layer conditions as relevant factors in the spatial and temporal variation of pollution concentrations. Each dataset contributes different information. Monitoring stations provide direct measurements of $PM_{2.5}$, but only for a limited number of locations. Satellite observations cover a much larger area, although aerosol optical depth is not a direct measure of particle concentrations near the ground. Reanalysis data add information about the atmospheric conditions that affect the transport, dispersion and accumulation of pollutants. By relating these sources to one another, the model produces a more spatially complete estimate of $PM_{2.5}$ than could be obtained from the station network alone.

Aji, Purnama and Husna [5] use a similar combination of data sources to examine riverine plastic pollution in Kendal Regency, Indonesia. Their analysis brings together Sentinel-2 imagery, land-use and land-cover data, precipitation records and a remotely sensed Plastic Index. Higher index values were generally associated with settlements and industrial areas, while forests and plantations tended to show lower values. The relationship with rainfall was weaker and varied across the study area. These results suggest that the distribution of the index is more closely related to the location of human activities than to precipitation alone. The study is useful in this respect because it does not treat every spatial correlation as equally informative. Instead, it allows several possible explanations for the observed pattern to be considered together, including rainfall, land use, industrial activity and settlement distribution. The same principle applies to the $PM_{2.5}$ study, where pollution levels reflect not only the location of emission sources but also the atmospheric conditions governing their dispersion. In both cases, the value of the model lies not simply in reproducing a spatial pattern, but in showing which geographical factors are most closely associated with it.

2.3. Learning from environmental time series

Geographical processes develop over time as well as across space. Several studies in the volume therefore examine whether AI methods can identify temporal relationships, predict environmental variables and improve the selection of suitable modelling approaches.

Nurdiati and colleagues [3] analyse monthly fire hotspots in Kalimantan using climatic indicators that include rainfall, the duration of dry spells, the El Niño–Southern Oscillation and the Indian Ocean Dipole. They compare Gaussian Process Regression with Long Short-Term Memory networks. The LSTM model performs better on the test data, suggesting that it captures relationships extending across successive time steps more effectively. Neither approach, however, reproduces all of the observed changes. In particular, both models have difficulty accounting for a substantial reduction in fire hotspots that appears to be associated with socioeconomic intervention rather than climatic variation. This result indicates a limitation of the predictor set rather than only of the algorithms themselves. Fire occurrence depends on rainfall and large-scale climate variability, but it is also influenced by land management, agricultural practices, enforcement, market conditions and policy measures. A model based exclusively on environmental variables can therefore represent only part of the process. The remaining prediction errors may provide an indication of where important human influences have not been included.

Jumadi and colleagues [6] compare a wider range of machine-learning and deep-learning methods for rainfall prediction in the Bengawan Solo River Basin. Nine algorithms are tested using a long CHIRPS rainfall record sampled at several hundred grid points. Extreme Gradient Boosting and a Gated Recurrent Unit model produce the most stable results, whereas some of the more elaborate deep-learning architectures show larger errors and stronger variation between months. The comparison shows that increasing model complexity does not necessarily improve rainfall prediction. Performance also depends on the length and consistency of the time series, the temporal resolution of the data, spatial differences within the basin and the way the models are trained and evaluated. For this reason, the comparison of algorithms is not simply a technical preparation for the analysis. It helps establish which modelling assumptions are best suited to the available data.

Taken together, the two studies also point to different sources of predictive uncertainty. In rainfall modelling, errors may arise because an algorithm does not capture the temporal structure of the data sufficiently well. In the fire-hotspot study, part of the uncertainty results from the absence of socioeconomic variables that could explain changes not driven by climate. A more complex temporal model may improve the representation of dependencies within a time series, but it cannot account for processes that are missing from the input data.

2.4. From prediction to geographical scenarios

Prediction in geographical research does not always refer to estimating a single value for the following day or month. It may also involve constructing spatial scenarios that show how a landscape could develop if recently observed patterns and relationships continue.

Triyatno and colleagues [2] apply this approach in the Kambang watershed of West Sumatra. They combine supervised land-cover classification with Cellular Automata–Markov Chain modelling and Random Forest estimation of landslide potential. Land-use and land-cover patterns are projected for 2029 and 2039, after which the simulated changes are used to assess possible shifts in landslide susceptibility. Particular attention is given to the expected reduction in primary forest and the expansion of other land uses. The resulting maps indicate both the projected extent of highly susceptible terrain and the areas where this increase is most likely to occur.

Sideng, Haris and Lamada [4] use a related modelling framework in Enrekang Regency, where an artificial neural network is combined with cellular automata to simulate land-use change and examine its relationship with flood risk. The neural network estimates the association between land-use transitions and their driving factors, while the cellular component accounts for the influence of neighbouring cells on the location of future change. Their results suggest that built-up areas may continue to expand into zones already exposed to flooding. The relevance of the simulation lies less in predicting the exact future state of individual pixels than in identifying where current development tendencies may create additional risk. Cellular automata are useful here because land-use change is spatially dependent. Development often follows existing settlements and roads, while slope, elevation and neighbouring land uses may either encourage or restrict expansion. The model therefore provides a conditional scenario: it shows how the landscape might develop if the relationships identified during the calibration period remain broadly similar.

Onphua, Teejuntuk and Na Takuathung [12] examine how the choice of calibration period affects this type of projection. In the Loe Tor Royal Project area, they compare several land-cover classification algorithms and test CA–Markov models based on three-, five- and ten-year intervals. The model calibrated over the shortest period produces the most reliable results, whereas the ten-year interval performs less well. The authors relate this difference to changes in transition probabilities over time and to spectral inconsistencies between images acquired by different sensors. Their findings show that a longer historical record is not necessarily more suitable for model calibration. A long interval may combine periods with different planning policies, technologies, land-management practices or data characteristics, even though the model treats them as parts of the same process. Selecting a calibration period therefore requires attention not only to the number of available years, but also to whether the observations represent a reasonably consistent geographical and institutional context.

2.5. Connecting data-driven models with physical knowledge

The volume also includes studies in which data-driven models are combined with physical constraints, process knowledge or variables designed to represent geographical conditions more directly. In these cases, AI is used as one component of the analysis rather than as an independent source of explanation.

Katiba, Belmajdoub and Minaoui [7] propose a hybrid framework in which a BiLSTM–MLP network predicts precipitation, while a reinforcement-learning controller tests small changes in selected atmospheric variables. The search is restricted by stability and energy constraints so that the simulated combinations remain physically plausible. The framework is presented as a tool for exploring the sensitivity of precipitation to changes in atmospheric conditions and for informing water-management analysis, not as a system for modifying weather in practice. The distinction is important because statistical associations between rainfall, pressure, humidity and wind do not imply that these variables can be controlled independently in the atmosphere. The main contribution of the study lies in the way it combines prediction with a constrained exploration of possible atmospheric states. Rather than allowing the controller to search without limits, the model excludes solutions that would conflict with basic physical conditions.

Nguyen and colleagues [11] incorporate physical information at an earlier stage of the modelling process. Their study develops a lithological weakness index from shear-strength data and uses it as an explanatory variable in GIS-based landslide-susceptibility modelling along mountainous road corridors in northern Vietnam. This provides a more direct representation of material resistance than a conventional classification based only on broad geological units. The study shows that model performance depends partly on how well the input variables reflect the processes being investigated. In landslide analysis, replacing a geological category with an indicator of mechanical weakness may improve the physical meaning of the model even when the underlying algorithm remains unchanged. The same consideration applies to other commonly used predictors, including slope, rainfall and land cover. These variables are not direct representations of environmental processes, but simplified measures constructed from particular datasets and assumptions. Their definition therefore requires geographical and physical judgement before they are introduced into the model.

3. WHAT GEOGRAPHY CONTRIBUTES TO AI

The studies discussed in the previous section show how AI can transform satellite images, time series and field measurements into flood maps, pollution estimates, rainfall forecasts, biomass inventories and land-use scenarios. Yet these outputs do not derive their geographical meaning from the algorithms alone. Before a model can recognise, estimate or predict anything, the research must define what constitutes the phenomenon, at what scale it should be observed, which relationships matter and how the result will be interpreted.

Geography therefore contributes more than spatial data to AI. It provides concepts for distinguishing meaningful entities, methods for organising heterogeneous evidence, an understanding of scale and spatial dependence, and knowledge of the environmental and social processes behind observed patterns. It also connects model outputs to places, institutions and decisions. Without this framework, an AI system may produce an accurate classification or prediction whose geographical significance remains uncertain (Janowicz et al., 2020; Wang et al., 2024; Mai et al., 2025).

3.1. Defining meaningful geographical entities

AI models require target categories, but geographical entities are rarely given in an unambiguous form. A flood boundary, burned area, land-use class, hazardous slope or urban neighbourhood must first be defined through observational rules and disciplinary knowledge. These definitions determine what the model learns.

Flood mapping illustrates the problem. Radar backscatter can help distinguish inundated surfaces, but low values may also represent permanent water, smooth artificial surfaces or radar shadow. The models used by Intarat, Kuankhamnuan and Jangsawang [1] can segment likely floodwater, but the resulting class becomes geographically meaningful only when it is interpreted in relation to floodplain morphology, drainage, land cover and the timing of image acquisition.

The preparation of reference data raises similar questions in burned-area mapping. Suppawimut and Samphutthanont [10] use the differenced Normalized Burn Ratio to create training masks for a deep-learning model. These masks are not neutral observations. Their meaning depends on the selected spectral threshold, the dates compared and the distinction between agricultural fire, harvest, bare soil and other forms of surface change. Geography contributes to AI here by defining the environmental process represented by the label.

Some geographical entities are multidimensional rather than directly visible. A geomorphosite, for example, is not simply a terrain form that can be recognised from elevation data. Irimuş et al.

evaluate volcanic landforms through scientific, scenic, cultural-historical, ecological and socioeconomic criteria (Irimuș et al., 2015). Computer vision might identify volcanic cones or craters, but it cannot derive the full significance of a site from morphology alone. The geographical classification combines observable form with scientific interpretation and place-based value.

Comparable issues arise in human geography. Accessibility may refer to proximity, travel time, available opportunities or the capacity of different population groups to reach them. Each definition measures something different and may lead to different planning conclusions (Geurs & van Wee, 2004). Health exposure is equally dependent on definition. A temperature value assigned to a residence, a neighbourhood average and the conditions experienced during daily mobility are not interchangeable measures (Clark et al., 2025).

The target variable is therefore already a geographical argument. AI can improve the consistency or speed of classification, but geography determines what the categories represent and which distinctions should remain visible.

3.2. Representing scale, connectivity and temporal structure

Geographical processes do not occur independently at individual pixels or observation points. They develop within neighbourhoods, catchments, transport networks, routes and regions. The spatial unit used in a model influence both the pattern detected and the conclusion drawn from it. Land-use change provides a clear example. Cellular automata are useful because the probability of transition depends partly on neighbouring cells. Development tends to follow existing settlements and roads, while slope, protected areas and surrounding land uses may constrain its direction. The studies by Triyatno et al. [2], Sideng et al. [4] and Onphua et al. [12] combine classification or machine-learning methods with CA–Markov approaches precisely because land-use transitions are spatially connected.

The choice of temporal interval matters as well. Onphua et al. [12] find that the shortest calibration period produces the most reliable projection. A longer record is not automatically better when it combines periods governed by different policies, land-management practices, technologies or sensor characteristics. Geography contributes by asking whether observations belong to the same underlying spatial and institutional process before they are treated as one training dataset.

Hydrological modelling faces a similar problem. Catchments differ in terrain, soils, vegetation, storage and river-network organisation. Models trained across many basins can identify shared rainfall–runoff relationships and transfer some information to poorly monitored regions (Kratzert et al., 2019). The catchment cannot, however, be treated as an arbitrary collection of grid cells. Flow connectivity, upstream–downstream relations and water balance remain part of the structure of the problem (Nearing et al., 2021).

Regional variation is equally important. Blöschl et al. found that river floods have increased in some parts of Europe and decreased in others, reflecting different combinations of precipitation, snow processes and catchment conditions (Blöschl et al., 2019). A model fitted to the continent as a whole could obscure these contrasting regional processes even if its overall statistical performance were acceptable.

Temporal relations also require geographical and process-based interpretation. Gasparrini et al. show that temperature may affect mortality over several subsequent days and that the exposure–response relationship differs between locations (Gasparrini et al., 2015). The relevant observation is therefore not simply the temperature and mortality recorded on the same date. Lag structure, seasonality and local adaptation form part of the phenomenon being modelled.

Route-level risk offers a smaller-scale illustration. Tourist-trail conditions depend on the sequence of terrain segments, exposure, weather, communication coverage and the characteristics of the user. A trail cannot be represented adequately by one average slope or a single hazard value (Magyari-Sáska, 2014). Its meaning comes from the connection between segments and from changing conditions along the route.

Geography thus contributes the units and relations through which AI interprets spatial data. It indicates when observations should be treated as neighbours, parts of a network, stages in a sequence or members of distinct regions.

3.3. Organising heterogeneous geographical evidence

Many geographical questions cannot be answered from one data source. The $PM_{2.5}$ study by Thabthong et al. [8], for example, combines monitoring stations, satellite aerosol optical depth and meteorological reanalysis. Each source observes a different aspect of atmospheric pollution. Ground stations measure near-surface concentrations directly but provide limited spatial coverage. Satellite data offer broader coverage but do not measure ground-level $PM_{2.5}$ directly. Reanalysis adds information on the atmospheric conditions controlling transport and dispersion.

The model can relate these sources statistically, but geographical reasoning is needed to understand why they can be combined and where their meanings differ. A continuous prediction surface should not conceal the fact that some values are measured while others are inferred.

Aji, Purnama and Husna [5] similarly combine a remotely sensed plastic index, rainfall and land-use information. The weaker relationship with precipitation and stronger association with settlements and industrial areas suggest that the spatial pattern cannot be explained by hydrology alone. Land use, waste generation and human activity offer competing or complementary interpretations. AI can rank variables or identify interactions, but the explanation requires knowledge of how pollutants enter and move through the river system.

Heterogeneity is even greater when numerical and documentary evidence are combined. Historical GIS links maps, archival texts, place names and changing administrative boundaries (Gregory & Healey, 2007). Natural-language processing may extract names and dates, but older spellings, transliterations and historical changes in place identity require contextual interpretation.

Multi-criteria evaluation brings together another mixture of evidence. In developing a geopark-attractiveness model, Nyulas et al. reviewed 52 assessment methods and 862 variables before selecting 33 attributes concerning natural and cultural heritage, accessibility, infrastructure, scientific value, safety and other dimensions (Nyulas et al., 2024). This reduction is not merely a technical feature-selection problem. It reflects the purpose of geopark evaluation, institutional requirements and decisions about which values should be represented.

Similar considerations apply to ecosystem-service assessment. Land-cover classes may provide a convenient basis for expert scoring, but the same class can support different services depending on soils, climate, management and landscape context (Burkhard et al., 2009). Machine learning can test these generalisations against observations, yet the services themselves remain defined through ecological knowledge and social priorities.

Geospatial knowledge graphs offer one way to preserve these distinctions. They can represent settlements, geosites, roads, protected areas, documents and organisations as different kinds of entities connected by explicit relationships (Bordogna & Fugazza, 2023; Mai et al., 2022). Their value lies not simply in storing more information, but in retaining the meaning and provenance of the links among sources.

Geography contributes to data integration by determining which forms of evidence are commensurable, which differences must be retained and how uncertainty should be carried from the original observations into the model output.

3.4. Keeping environmental and social processes visible

Predictive performance does not necessarily provide an explanation. A model can reproduce an observed pattern without identifying the process that generated it. Geographical knowledge is therefore needed both when predictors are selected and when the remaining errors are interpreted.

The fire-hotspot study by Nurdiati et al. [3] provides a useful example. Climate variables explain part of the temporal variation, and the LSTM model performs better than Gaussian Process Regression. Neither model fully captures a marked decline in hotspot occurrence that appears to be related to socioeconomic intervention. The error is informative because it suggests that the predictor set omits land-management, regulatory or behavioural changes. Increasing network complexity would not solve the absence of relevant variables.

Rainfall modelling leads to a related conclusion. Jumadi et al. [6] compare nine algorithms and find that some relatively compact models perform more consistently than more elaborate deep-learning architectures. The choice of method must correspond to the length, temporal resolution and spatial structure of the available record. Complexity alone does not supply missing process information.

The physical meaning of predictors also matters. Nguyen et al. [11] derive a lithological weakness index from shear-strength data rather than relying only on broad geological classes. The improvement lies partly in representing material resistance more directly. The example shows that a geographically and physically meaningful variable may contribute more than a change of algorithm.

Katiba, Belmajdoub and Minaoui [7] take this further by restricting a reinforcement-learning search through atmospheric stability and energy constraints. The controller is not allowed to explore every statistically advantageous combination of variables. Its solutions must remain compatible with basic physical conditions.

Theory-guided and hybrid approaches generalise this principle. Conservation laws, process-model outputs and known environmental relationships can be incorporated into machine-learning systems to reduce physically implausible predictions (Karpatne et al., 2017; Reichstein et al., 2019; Irrgang et al., 2021). Such models recognise that the available data may not contain enough examples to teach every relevant constraint.

Social processes require similar caution, although they cannot usually be represented through physical equations. Satellite imagery may predict economic conditions from buildings, roads and cultivated land, but it cannot directly observe tenure, household relations or informal employment (Jean et al., 2016; Yeh et al., 2020). A transport model may learn low demand in a poorly served neighbourhood without recognising that the lack of service itself suppresses travel.

Geography helps distinguish prediction from explanation. It asks whether a model has identified a process, a proxy or merely a stable correlation, and whether the relationship is likely to persist when environmental or institutional conditions change.

3.5. Connecting model outputs to geographical decisions

The output of an AI model becomes useful only when it can inform a geographical decision. A flood mask must be related to settlements, roads, agricultural land or evacuation routes. A pollution surface must be interpreted in relation to population exposure. A land-use scenario must be evaluated against planning objectives and environmental constraints.

These decisions introduce values that cannot be inferred from accuracy alone. A highly accurate accessibility model cannot determine whether the existing distribution of services is fair. A tourism recommender trained on previous visits may reproduce the concentration of visitors in already popular places. A geopark assessment must balance conservation, accessibility, education and local development. The algorithm can compare alternatives, but it cannot establish the priorities on its own.

Scenario models require similar interpretation. The land-use projections in the volume do not describe a predetermined future. They show what may occur if the relationships observed during the calibration period continue. Their planning value lies in identifying where present tendencies could increase flood or landslide exposure, not in claiming certainty about individual future pixels.

Operational warning systems add questions of responsibility. Flood forecasts, heat warnings and trail-risk assessments should indicate where observations are sparse, when the data were updated and how certain the prediction is. A model may support a decision, but formal responsibility remains with the relevant authorities and professional organisations.

Cartography also plays a role in this transition. Model results must be presented in a form that allows users to distinguish predicted values, uncertainty, observed data and areas of extrapolation. Automated mapping can improve efficiency, but map design remains an interpretive act involving classification, generalisation and visual emphasis (Usery et al., 2022).

The geographical contribution continues after the model has produced its prediction. It connects the result to affected places and populations, evaluates whether the output is suitable for its intended purpose and makes the assumptions and uncertainties visible.

3.6. Geography as part of the model, not merely its application area

The twelve studies in this volume demonstrate the breadth of AI applications in geographical research. They also show that the most successful models depend on decisions made outside the algorithm: how reference classes are defined, how spatial units are connected, which temporal period is representative, which variables reflect the process and how outputs are linked to action.

Geography contributes these decisions. It defines meaningful entities, represents scale and connectivity, organises heterogeneous evidence, supplies environmental and social interpretation, and relates predictions to institutional and public objectives. Field inventories, epidemiological models, accessibility frameworks, heritage assessments and physical process knowledge are not preliminary forms of AI. They are sources of the structure that AI needs.

This perspective also changes how model improvement should be understood. A deeper architecture or larger dataset may raise predictive accuracy, but improvement may also come from a better geographical unit, a more meaningful predictor, an appropriate temporal lag or a clearer distinction between observation and inference. In many cases, these choices have a greater effect on the usefulness of the result than the selection of the algorithm.

GeoAI is therefore strongest when geography remains present throughout the research process: before modelling, when the problem and data are defined; within the model, when spatial relations and process constraints are represented; and after modelling, when predictions are interpreted, mapped and used. Geography does not simply supply locations for artificial intelligence. It provides the concepts through which spatial predictions become knowledge.

4. CLOSING REFLECTIONS

The value of an AI model in geography depends not only on predictive accuracy, but also on whether it is tested under realistic spatial conditions. Nearby observations often share similar environmental and social characteristics, so randomly dividing them between training and test datasets may produce overly optimistic results. In such cases, the model is largely interpolating within familiar surroundings rather than predicting in genuinely new areas. Spatially informed cross-validation provides a more reliable assessment, although its design must reflect the intended application, scale and sampling structure (Wang et al., 2023).

Transferability must likewise be demonstrated rather than assumed. A model developed for one city, watershed, mountain region or tourism destination learns relationships shaped by local climate, morphology, land use, institutions and data quality. Transfer learning may reduce the amount of new training data required, but it cannot replace local calibration and independent validation (Ma et al., 2024). A transferable methodology is therefore not the same as a fitted model that can be applied unchanged in every region.

Predictions must also be understandable to their intended users. A geomorphologist may need to know which terrain properties produced a classification, a planner which criteria altered a suitability score, and a hiker whether a warning reflects steep slopes, snow or an approaching storm. General feature-importance rankings are often insufficient because they may neither explain individual predictions nor establish causal relationships. Explanations should therefore combine model diagnostics with disciplinary knowledge, uncertainty and the institutional context in which the result will be used (Gevaert, 2022).

Uncertainty should be communicated spatially. A single accuracy value may conceal poor performance in particular regions, seasons, terrain types or land-cover classes. Confidence surfaces, prediction intervals and maps showing distance from the training domain can reveal where results are well supported and where they depend mainly on extrapolation. This is particularly important in health, safety and environmental management, where the consequences of errors vary geographically.

The choice of predictors, target variables and evaluation criteria remains a geographical decision. AI can optimise a model after the research problem has been defined, but it cannot independently decide whether tourism development should prioritise accessibility or conservation, whether a route is acceptably safe, or whether a landform has cultural significance. Such questions require scientific interpretation, local knowledge, institutional objectives and public values.

The wider use of AI in geography can be understood as a progression from recognising spatial entities, through estimating and comparing their characteristics, to anticipating change and supporting decisions. Across these stages, established geographical frameworks define the entities, relationships, scales and evaluation criteria that AI systems must respect.

AI is therefore most useful not when it replaces GIS, field surveys, statistical analysis, physical modelling or expert interpretation, but when it extends them. It can process more observations, update assessments more frequently and compare alternatives that would be difficult to evaluate manually. The studies considered here also show that greater model complexity is not automatically beneficial: deeper architectures, longer calibration periods or more elaborate methods may perform less reliably than simpler alternatives.

The future of AI in geography will depend on careful integration with geographical knowledge. Models must be grounded in place, validated across meaningful spatial and temporal separations, interpreted through environmental and social processes, and accompanied by visible uncertainty. The objective is not simply to produce more predictions, but to create spatial knowledge that can be examined, challenged and used responsibly.

REFERENCES

Blöschl, G., Hall, J., Viglione, A., et al. (2019). Changing climate both increases and decreases European river floods. *Nature, 573*, 108–111. https://doi.org/10.1038/s41586-019-1495-6

Bordogna, G., & Fugazza, C. (2023). Artificial intelligence for multisource geospatial information. *ISPRS International Journal of Geo-Information, 12*(1), 10. https://doi.org/10.3390/ijgi12010010

Burkhard, B., Kroll, F., Müller, F., & Windhorst, W. (2009). Landscapes' capacities to provide ecosystem services: A concept for land-cover-based assessments. *Landscape Online, 15*, 1–22. https://doi.org/10.3097/LO.200915

Clark, L. P., Zilber, D., Schmitt, C., Fargo, D. C., Reif, D. M., Motsinger-Reif, A. A., & Messier, K. P. (2025). A review of geospatial exposure models and approaches for health data integration. *Journal of Exposure Science & Environmental Epidemiology, 35*(2), 131–148. https://doi.org/10.1038/s41370-024-00712-8

Gasparrini, A., Guo, Y., Hashizume, M., et al. (2015). Mortality risk attributable to high and low ambient temperature: A multicountry observational study. *The Lancet, 386*(9991), 369–375. https://doi.org/10.1016/S0140-6736(14)62114-0

Gevaert, C. M. (2022). Explainable AI for Earth observation: A review including societal and regulatory perspectives. *International Journal of Applied Earth Observation and Geoinformation, 112*, 102869. https://doi.org/10.1016/j.jag.2022.102869

Geurs, K. T., & van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: Review and research directions. *Journal of Transport Geography, 12*(2), 127–140. https://doi.org/10.1016/j.jtrangeo.2003.10.005

Gregory, I. N., & Healey, R. G. (2007). Historical GIS: Structuring, mapping and analysing geographies of the past. *Progress in Human Geography, 31*(5), 638–653. https://doi.org/10.1177/0309132507081495

Irrgang, C., Boers, N., Sonnewald, M., Barnes, E. A., Kadow, C., Staneva, J., & Saynisch-Wagner, J. (2021). Towards neural Earth system modelling by integrating artificial intelligence in Earth system science. *Nature Machine Intelligence, 3*, 667–674. https://doi.org/10.1038/s42256-021-00374-3

Irimuș, I. A., Bálint-Bálint, L., Dombay, Ș., Crișan, H. F., & Magyari-Sáska, Z. (2015). Classification and evaluation criteria for volcanic geomorphosites in Harghita Mountains. In *Proceedings of the 15th International Multidisciplinary Scientific GeoConference SGEM 2015* (Book 1, Vol. 1, pp. 77–84). https://doi.org/10.5593/SGEM2015/B11/S1.010

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, 34*(4), 625–636. https://doi.org/10.1080/13658816.2019.1684500

Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016). Combining satellite imagery and machine learning to predict poverty. *Science, 353*(6301), 790–794. https://doi.org/10.1126/science.aaf7894

Karpatne, A., Atluri, G., Faghmous, J. H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N., & Kumar, V. (2017). Theory-guided data science: A new paradigm for scientific discovery from data. *IEEE Transactions on Knowledge and Data Engineering, 29*(10), 2318–2331. https://doi.org/10.1109/TKDE.2017.2720168

Kratzert, F., Klotz, D., Shalev, G., Klambauer, G., Hochreiter, S., & Nearing, G. (2019). Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets. *Hydrology and Earth System Sciences, 23*(12), 5089–5110. https://doi.org/10.5194/hess-23-5089-2019

Ma, Y., Chen, S., Ermon, S., & Lobell, D. B. (2024). Transfer learning in environmental remote sensing. *Remote Sensing of Environment, 301*, 113924. https://doi.org/10.1016/j.rse.2023.113924

Magyari-Sáska, Z. (2014). Quantifying threats along tourist trails: An initial approach. *Geographia Technica, 9*(1), 78–86.

Mai, G., Hu, Y., Gao, S., Cai, L., Martins, B., Scholz, J., Gao, J., & Janowicz, K. (2022). Symbolic and subsymbolic GeoAI: Geospatial knowledge graphs and spatially explicit machine learning. *Transactions in GIS, 26*(8), 3118–3124. https://doi.org/10.1111/tgis.13012

Mai, G., Xie, Y., Jia, X., Lao, N., Rao, J., Zhu, Q., Liu, Z., Chiang, Y.-Y., & Jiao, J. (2025). Towards the next generation of geospatial artificial intelligence. *International Journal of Applied Earth Observation and Geoinformation, 136*, 104368. https://doi.org/10.1016/j.jag.2025.104368

Nearing, G. S., Kratzert, F., Sampson, A. K., Pelissier, C. S., Klotz, D., Frame, J. M., Prieto, C., & Gupta, H. V. (2021). What role does hydrological science play in the age of machine learning? *Water Resources Research, 57*(3), e2020WR028091. https://doi.org/10.1029/2020WR028091

Nyulas, J., Dezsi, Ș., Haidu, I., Magyari-Sáska, Z., & Niță, A. (2024). Attractiveness assessment model for evaluating an area for a potential geopark—Case study: Hațeg UNESCO Global Geopark (Romania). *Land, 13*(2), 148. https://doi.org/10.3390/land13020148

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature, 566*, 195–204. https://doi.org/10.1038/s41586-019-0912-1

Usery, E. L., Arundel, S. T., Shavers, E. J., Stanislawski, L. V., Thiem, P. T., & Varanka, D. E. (2022). GeoAI in the US Geological Survey for topographic mapping. *Transactions in GIS, 26*(1), 25–40. https://doi.org/10.1111/tgis.12830

Wang, S., Huang, X., Liu, P., Zhang, M., Biljecki, F., Hu, T., Fu, X., Liu, L., Liu, X., Wang, R., Huang, Y., Yan, J., Jiang, J., Chukwu, M., Reza Naghedi, S., Hemmati, M., Shao, Y., Jia, N., Xiao, Z., … Bao, S. (2024). Mapping the landscape and roadmap of geospatial artificial intelligence (GeoAI) in quantitative human geography: An extensive systematic review. *International Journal of Applied Earth Observation and Geoinformation, 128*, 103734. https://doi.org/10.1016/j.jag.2024.103734

Wang, Y., Khodadadzadeh, M., & Zurita-Milla, R. (2023). Spatial+: A new cross-validation method to evaluate geospatial machine-learning models. *International Journal of Applied Earth Observation and Geoinformation, 121*, 103364. https://doi.org/10.1016/j.jag.2023.103364

Yeh, C., Perez, A., Driscoll, A., Azzari, G., Tang, Z., Lobell, D. B., Ermon, S., & Burke, M. (2020). Using publicly available satellite imagery and deep learning to understand economic well-being in Africa. *Nature Communications, 11*, 2583. https://doi.org/10.1038/s41467-020-16185-w