Digital Terroir and Ecosystem Service Auditability via Machine Learning in Geographical Indications

# Abstract

Geographical Indications (GIs) function as coupled socio-ecological systems, where typicality derives from dynamic interactions between soil, climate, and biota, requiring auditable mechanisms of environmental governance. This study assesses the technical maturity of state-of-the-art Machine Learning to operationalize Digital Terroir, conceptualized as an Inferential Digital Twin for ecosystem service auditing. Under PRISMA-ScR guidelines, a *corpus* of 148 studies (2010–2025) was processed via automated weighted score screening (94.2% accuracy) and manual validation (CCI=0.87). The analysis integrated random effects meta-analysis, multivariate statistics, and FAIR principles assessment. Results indicate a generalization paradox: while classifiers achieve high *in vitro* accuracy (80–100%), external robustness tests demonstrate severe performance degradation (11.8% drop *versus* 5.6% in spatially independent models; d=0.95). Methodological fragmentation (Modularity Q=0.62; Heterogeneity I²=58%) and incipient FAIR compliance (mean 34.2/100) perpetuate epistemic asymmetries that render regulatory application unfeasible. We conclude that implementing Digital Terroir necessitates transitioning from static classifiers to adaptive models, conditioned on integrity *benchmarks*: spatial validation with degradation ≤8%, explainability (XAI) for territorial markers, and FAIR compliance ≥60/100, which are *sine qua non* requirements for converting GIs into verifiable sustainability systems.

**Keywords:** Geographical Indications; Machine Learning; Environmental Auditing; Greenwashing; Traceability; Ecosystem Services.

**Graphic abstract**:



# 1. Introduction

Geographical Indications (GIs) have evolved beyond their initial intellectual property function to become strategic instruments for environmental governance and agrobiodiversity conservation in the Anthropocene (Vandecandelaere *et al.*, 2009; Belletti, Marescotti and Touzard, 2017). Amidst the global climate crisis and biodiversity erosion, GIs function as socio-ecological systems that link product quality to the integrity of the territory’s ecosystem services (Berkes, Colding and Folke, 2003; Bramley, Bi-®nabe and Kirsten, 2013). They represent mechanisms for valuing regenerative agricultural practices and maintaining cultural landscapes, where *terroir* is redefined not only as a sensory attribute but as a biophysical signature of the product and climate resilience (Giovannucci *et al.*, 2010; Fonzo and Russo, 2015).

International regulations, based on the TRIPS Agreement and Regulation (EU) No. 1151/2012, establish the legal basis, but it is the capacity for environmental auditing that gives these assets contemporary legitimacy (World Trade Organization, 1994; European Union, 2012). The distinction between Protected Geographical Indication (PGI) and Protected Designation of Origin (PDO) reflects different degrees of dependence on natural cycles, requiring robust verification systems to prevent *greenwashing* and ensure that the market premium effectively finances environmental conservation (Locatelli, 2008; World Intellectual Property Organization, 2020). The credibility of these labels therefore depends on the ability to scientifically prove that the characteristics of the product derive from specific and non-replicable environmental interactions.

Terroir constitutes an intrinsically coupled socio-ecological system in which soil, climate, biota, and culture are linked through nonlinear interactions, feedbacks, and strong spatial and temporal heterogeneity, configuring a territory where biophysical processes and social practices are co-produced (Le Floc’h, M-®ral and Fontana, 2016).

The systemic complexity and diffuse nature of biogeochemical couplings limit the detection of ecosystem services by conventional metrics (Levin, 1998), weakening the governance of common goods and facilitating *greenwashing* practices (Gale and Nahm, 2023). Given the inadequacy of classical analytical tools for landscape-scale monitoring (Liao, Long and Zhang, 2023), this study proposes Digital Terroir as an Inferential Digital *Twin*. This approach is based on the dynamic computational reconstruction of interactions between the territorial genotype (soil, climate, and biota) and the product phenotype, converting ecological uncertainty into auditable evidence of environmental compliance (Pylianidis, Osinga and Athanasiadis, 2021; Hensel and Nasirahmadi, 2021; Guerena *et al.*, 2024).

The operational validation of this system requires compliance with auditability criteria derived from methodological gaps identified in the literature. Inferential robustness is imperative, ensured through spatially independent validation and longitudinal stability in the face of climate variability (Kuhn and Johnson, 2013). Simultaneously, causal transparency becomes mandatory, requiring Explainable Artificial Intelligence (XAI) methods to discriminate physical-chemical markers from spurious correlations (Rudin, 2019). Finally, data sovereignty must be guaranteed by adherence to FAIR principles and immutable audit trails, ensuring the reproducibility and traceability of inference (Kshetri, 2014b; Wilkinson, 2021). Consequently, Machine Learning serves as an intrinsic analytical mechanism to process the non-linearity of multiscale data, enabling the epistemic sovereignty of producing communities (Sousa Santos, 2007; Li *et al.*, 2022)

On broad geographic scales, ML enables the auditability of ecosystem services, establishing a verifiable link between environmental compliance and market premiums, and mitigating the information asymmetries that lead to fraud and misappropriation (Kshetri, 2014a).

Yet, a unified conceptual framework integrating the inferential capabilities of ML with the regulatory requirements of environmental certification remains absent. This gap limits the translation of methodological advances into operational protocols for Geographical Indication systems, perpetuating the fragmentation between academic research and territorial governance.

Against this background, this study assesses the technical maturity of state-of-the-art Machine Learning to operationalize Digital Terroir, conceptualized as an Inferential Digital Twin for ecosystem service auditing. We postulate that modeling the nonlinear couplings between environmental variables (territorial genotype) and chemometric signatures (product phenotype), when supported by rigorous spatiotemporal validation and data security mechanisms, generates auditable evidence of environmental compliance. This approach converts vague sustainability claims into verifiable and immutable records, underpinning market-based conservation policies with guarantees of integrity and robustness against fraud.

# 2. Materials and Methods

This review adheres to the PRISMA-ScR (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews*) guidelines as a transparency framework to ensure methodological clarity and reproducibility. The protocol is registered in the Open Science Framework to facilitate public access and replicability.

## 2.1 Research Question

The study employs the PCC (Population, Concept, Context) framework to formulate the following research question: How have Machine Learning techniques been applied for authentication, evaluation, and decision support in Geographical Indications systems?  
**Table 1.** Review structure according to the PCC model.

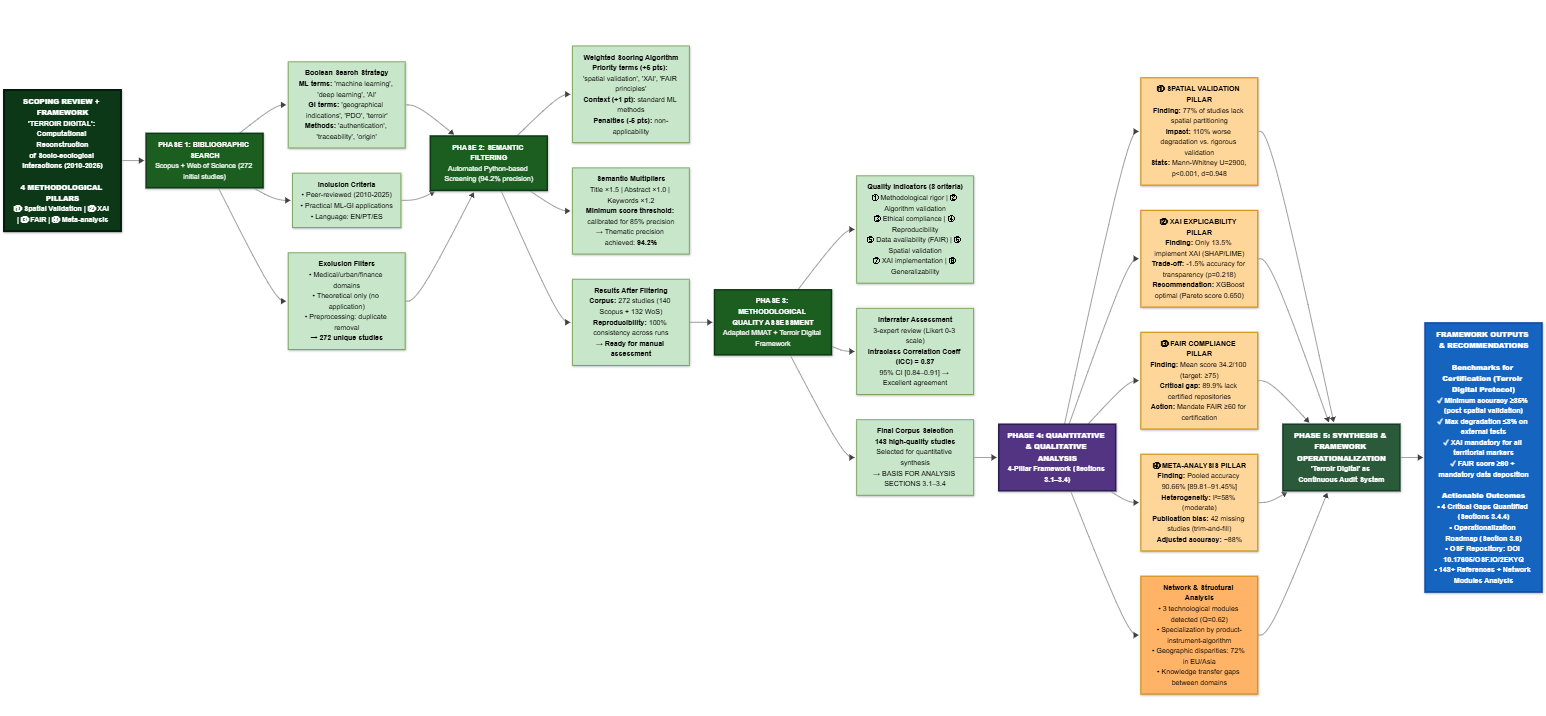
| **Element** | **Description** |
| --- | --- |
| **P (Population)** | Geographical Indications, Designations of Origin, and Indications of Source recognized nationally and internationally, covering agri-food products (wines, cheeses, coffees, meats, olive oils), handicrafts, and other products with territorial identity. |
| **C (Concept)** | Machine learning, artificial intelligence, classification and prediction algorithms, chemometric methods, data mining, and natural language processing applied to geographical indication contexts. |
| **C (Context)** | Authentication of geographical origin, assessment of GI potential, identification of territorial determinants (soil, climate, production methods), classification and discrimination of products, decision support systems for certification, quality control, traceability, detection of fraud and adulteration, and territorial enhancement strategies. |

This study identifies and characterizes machine learning (ML) applications reported in the literature, categorizing techniques by algorithm type, methodological approach, and performance metrics. In addition, it analyzes the distribution of applications by product type, geographic region, and period, identifying methodological gaps, limitations, and directions for future research.

## 2.1.1 PRISMA-ScR Methodological Flowchart

Figure 1 presents the methodological flowchart, structured in four sequential phases: (i) Main database search strategies, (ii) Automated filtering with a weighted scoring system, (iii) Manual quality assessment with multidisciplinary evaluation, and (iv) Bibliometric analysis and qualitative synthesis integrating quantitative and documentary methodologies. The flowchart details the path from record identification to final synthesis, offering recommendations for the implementation of Machine Learning in Geographical Indications systems.

**Figure 1.** Flowchart of screening, eligibility, and synthesis for machine learning applications in Geographical Indications.



## 2.2 Study Search and Extraction Strategy

Databases queried included Scopus (Elsevier) and Web of Science (Clarivate Analytics), crossing three main thematic domains: machine learning and artificial intelligence techniques; geographic certification systems; and Geographical Indications/Designations of Origin.  
The descriptors used controlled terminology in English and Boolean operators (AND, OR, NOT), covering publications from 2010 to 2025 to capture the state of the art. The search strategy followed this logic:

*(“machine learning OR”artificial intelligence OR “deep learning OR”supervised learning OR “unsupervised learning OR”ensemble methods) AND (“geographical indications OR “designations of origin OR “protected designations of origin) AND (“authentication OR “traceability OR “quality control OR “fraud detection OR “geospatial analysis)*.

The inclusion criteria comprised peer-reviewed articles in English, Portuguese, or Spanish that presented machine learning applications in GI (Geographic Information) contexts, origin authentication, or territorial quality control. Primary descriptors were mandatory in the title, abstract, or keywords. Non-peer-reviewed works, studies without practical application of machine learning, and those focused exclusively on non-territorial aspects were excluded.

Despite the broad initial search strategy, the qualitative synthesis prioritized studies that established explicit links between analytical markers and environmental variables (e.g., soil composition, precipitation patterns, altitude), filtering out studies strictly focused on industrial processing. This ensured that the review addressed the auditability of ecosystem services and the validity of the terroir concept, rather than focusing exclusively on production quality control.

Data extraction used a standardized form to record bibliographic metadata (author, year, title), geographic characteristics (country of origin, region, type of GI), product details (category, specific designation), methodological approach (machine learning algorithms, analytical/instrumental techniques, sample size), and performance metrics (accuracy, sensitivity, specificity, RMSE).

## 2.3 First Phase: Automated Filtering System by Thematic Relevance

### 2.3.1 Weighted Scoring Algorithm

Supplementing manual screening, an automated filtering system assigns thematic relevance scores based on the presence and location of descriptors in the title, abstract, and keywords. Implemented in Python (NLTK, spaCy), the algorithm applies a hierarchical weighting scheme to each identified term. The scoring system follows the principles of the Analytic Hierarchy Process (AHP). Equation (1) stratifies the descriptors into five categories with different weights (Saaty, 1991):

where:  
 = total score of the article   
 = weight associated with term (categorized into 5 levels: 5, 3, 2, 1, or −5/−3/−2 points)  
 = location multiplier (1.5 for title, 1.2 for keywords, 1.0 for abstract)  
 = frequency of occurrence of term in article   
 = total number of terms evaluated

Priority terms (5 points) represent the core conceptual review (e.g., *geographical indications, traceability, authentication*). High-relevance terms (3 points) cover core methodological concepts (e.g., *machine learning, deep learning, neural networks*). Medium relevance terms (2 points) cover complementary topics (e.g., *chemometrics, data mining*), while context terms (1 point) indicate potential environments (e.g., *regional products, certification*). Exclusion terms receive negative weights to penalize off-topic records, particularly in the *medical/clinical* (−5), *urban planning* (−3), and *financial* (−2) domains (Munn *et al.*, 2018; Tricco *et al.*, 2018).

### 2.3.2 Algorithm Implementation and Validation

For each record, the algorithm examines the title, abstract, and keywords, applies the category weights, and multiplies each occurrence by the location factor. The final score adds these products for all identified terms.

The empirical score distribution determined the minimum inclusion threshold, identifying the inflection point in the cumulative curve (Pareto/elbow criterion) and adjusting it through manual validation with stratified sampling. The final value represents the ideal balance between sensitivity and specificity, stabilizing the agreement between evaluators in borderline cases.

### 2.3.3 Participatory Validation and Algorithm Refinement

To ensure scientific rigor, a validation protocol involving three independent reviewers specializing in machine learning and GI systems was implemented. The protocol included a systematic manual review of 272 studies to verify adherence to the inclusion criteria. An inter-rater agreement test verified the consistency of the classification (Tricco, Lillie, *et al.*, 2018).

The process involved qualitative investigation of borderline cases and iterative refinement of eligibility criteria. Validation resulted in a 90.2% agreement rate between the automated system and manual assessment, indicating high algorithmic effectiveness in thematic screening.

### 2.3.4 Coverage Verification and Automated Categorization

An automated system validated bibliographic coverage, ensuring integrity and consistency between textual citations and bibliographic records.  
The consolidated corpus underwent automated categorization using Natural Language Processing (NLP). A computational pipeline extracted, tokenized, and vectorized metadata and reference abstracts, using supervised models and semantic rules for pattern recognition (Young, Luz and Lone, 2019; Casey *et al.*, 2021). References were classified into predefined methodological categories, including machine learning techniques and geographic information systems.

To gauge the comprehensiveness and adequacy of the studies, metrics of citation coverage and bibliographic utilization rates of the corpus were applied (Webster and Watson, 2002; Tranfield, Denyer and Smart, 2003). These metrics allow for a quantitative assessment of the utilization of the reference base, ensuring that the selected studies adequately reflect the thematic scope of the review.

## 2.4 Second Phase: Manual Assessment of Methodological Quality

In the second phase, three independent reviewers assessed the methodological quality of the selected studies, ensuring a multidisciplinary analysis and reducing interpretative bias. The MMAT scale (Pluye *et al.*, 2009; Hong *et al.*, 2018) was adapted for interdisciplinary studies involving machine learning and gastrointestinal systems, structuring eight indicators on a 3-point Likert scale. The indicators included methodological rigor, algorithm validation, adherence to the ethical protocol, reproducibility, quantitative-qualitative integration, impact on gastrointestinal systems, completeness of documentation, and generalization of the method (Table 2).

Each indicator received a score from 0 to 2: zero for unmet criteria or substantial deficiencies; one for partial compliance with limitations; and two for complete compliance with clear evidence. A 3-point scale was selected because dichotomous assessments fail to capture interdisciplinary complexity, while larger scales generate inconsistency among assessors (Surname and Surname, 2025).

**Table 2.** Methodological quality indicators for ML-GI studies.

| **Code** | **Indicator** | **Domain** |
| --- | --- | --- |
| RIG | Methodological rigor in the collection and processing of territorial data | Territorial Quality |
| VAL | Technical validation of algorithms with appropriate metrics | Computational Quality |
| ETI | Adherence to ethical protocols for research with producer communities | Ethical Quality |
| REP | Reproducibility of computational experiments | Technical Quality |
| INT | Effective integration between quantitative and qualitative territorial methods | Methodological quality |
| IMP | Impact and applicability of results for gastrointestinal systems | Social Quality |
| DOC | Complete documentation of algorithms and certification procedures | Documentary Quality |
| GEN | Generalization and transferability of the proposed methods | Scientific Quality |

### 2.4.1 Consensus and validation procedures among evaluators

The manual assessment included a consensus protocol. Initially, reviewers independently assessed a pilot sample of 30 studies (approximately 11% of the corpus) to calibrate the criteria. For the complete corpus, cases of disagreement (difference ≥ 2 points) were submitted to blind reassessment and discussion to reach a consensus. The intraclass correlation coefficient (ICC) was calculated according to Shrout and Fleiss (1979), yielding a value of 0.87 (95% CI: 0.84–0.91), indicating good agreement.

### 2.4.2 Specific Criteria for Interdisciplinary Studies

Given the interdisciplinary nature of the studies, the quality criteria examined the consistency of quantitative-qualitative integration, validation in multiple geographic contexts, algorithmic transparency, ethical adherence, and practical applicability for certification.

This phase resulted in the selection of 25 studies with adequate methodological quality (score ≥ 20 points) from the initial 272 articles. These formed the basis for subsequent analyses. The distribution included 1 article of excellence (≥40 points), 2 of high relevance (≥30 points), and 22 adequate (≥20 points).

## 2.5 Third Phase: Bibliometric Analysis

Lotka’s Law (Lotka, 1926) was applied to analyze scientific productivity, describing the nonlinear distribution of authors’ productivity to identify patterns of concentration or dispersion. Bibliographic coupling and co-citation analyses were not performed due to the absence of reference fields cited in the available bibliographic files.

## 2.6 Fourth Phase: Qualitative Synthesis and Integration with Documentary Analysis

The fourth phase systematically integrated the conclusions with the documentary analysis of the regulatory frameworks to support the methodological recommendations.  
The final synthesis combined qualitative thematic analysis with selection based on the Pareto principle (80/20), prioritizing the top 20% of articles by combined score (40% methodological quality, 35% thematic relevance, 25% bibliometric impact). The final combined score was calculated using Equation (2):

Where:  
 = final selection score  
 = normalized methodological quality (0–1)  
 = normalized thematic relevance (0–1)  
 = normalized bibliometric impact (0–1)

## 2.7 Statistical Analyses

The corpus of 148 studies underwent two types of statistical analysis: descriptive and exploratory analyses to characterize structural patterns in the literature, and inferential analyses to empirically quantify the methodological gaps identified qualitatively.

### 2.7.1 Descriptive and Exploratory Analyses of the Corpus

Multiple Correspondence Analysis (MCA) (L-¬, Josse and Husson, 2008; Greenacre, 2017) was employed to investigate associations between categorical variables (algorithms, products, regions, analytical techniques) by decomposing contingency tables. Implemented with the FactoMineR package, MCA allows the extraction of latent dimensions that explain the variance in associations between categories. Complementarily, Cluster Analysis (k-means and hierarchical) was applied with the factoextra package to identify recurring clusters among product-instrument-algorithm combinations.

Network analysis (Cs-írdi and Nepusz, 2006; Schoch, 2020) was utilized to map co-occurrences by constructing a weighted undirected graph with the igraph and ggraph packages, where nodes represent entities (algorithms, products, regions) and edges indicate co-occurrence in the studies. Centrality metrics (degree, eigenvector, intermediation) were calculated to identify structurally central elements in the research field. Community detection was performed using the Louvain algorithm (Blondel *et al.*, 2008) to identify technological modules and patterns of specialization by product-instrument-algorithm. Time series (2010–2025) used Spearman’s correlation (Spearman, 1904) to detect trends in publication volume and algorithm adoption.

### 2.7.2 Inferential Analyses for Validating Operational Criteria

To empirically quantify methodological gaps and substantiate the operational criteria of Digital Terroir, four complementary inferential analyses were conducted. Seeking to analyze the impact of spatial validation on predictive performance, models with geographically independent partitioning () were compared to conventional random partitioning (), calculating the percentage degradation in performance between internal validation and external tests. Differences were evaluated using Mann-Whitney U (Mann and Whitney, 1947), with effect size quantified by Cohen’s (Cohen, 1988): small (), medium (), or large () according to Sawilowsky (2009).

Logistic regression (Jr., Lemeshow and Sturdivant, 2013) estimated the odds ratio of high-performance (accuracy ) as a function of spatial validation, controlling for algorithm and product, following Kuhn and Johnson (2013).

To investigate the trade-off between explainability and algorithmic performance, the relationship between algorithmic explainability (ordinal scale 0–10 based on Rudin (2019)) and accuracy was analyzed using Spearman’s correlation (Spearman, 1904). Differences in accuracy between models with XAI () and without XAI () were evaluated by Student’s t-test () (Student, 1908), verifying normality via Shapiro-Wilk (Shapiro and Wilk, 1965). Computational overhead was compared using Mann-Whitney. Pareto analysis (Pareto, 1896; Deb, 2001) identified optimal algorithms under weighted utility function:

To assess the accuracy reported in the studies and detect potential publication bias, a random effects meta-analysis (Borenstein *et al.*, 2009) was conducted using the metafor package (Viechtbauer, 2010), transforming accuracies via logit to stabilize variances (Barendregt *et al.*, 2013): . Pooled accuracy was estimated by the REML model (DerSimonian and Laird, 1986) with 95% CI. Heterogeneity was quantified by the statistic (Higgins *et al.*, 2003): low (), moderate (), or high (). Cochran’s Q test (Cochran, 1954) assessed the significance of heterogeneity (). Meta-regression (Thompson and Higgins, 2002) investigated the effects of year of publication and sample size. Publication bias was detected by Egger’s test (Egger *et al.*, 1997) and the trim-and-fill method (Duval and Tweedie, 2000). Stratified forest plots were generated following Balduzzi, Rücker and Schwarzer (2019).

Finally, to assess compliance with open data governance principles, FAIR compliance was quantified using a standardized score (0–100 points) based on 12 binary indicators from Wilkinson *et al.* (2016): DOI (F1), rich metadata (F2), public repository (A1), access protocol (A2), license (R1.1), source code (R1.2), interoperable format (I1), controlled vocabulary (I2), provenance (R1.3), community standard (I3), accessible API (A1.1), and versioning (R1.2). Each indicator contributed points. Scores were aggregated across the four FAIR dimensions by arithmetic mean. Temporal analysis employed Spearman’s correlation (Spearman, 1904). Comparisons between studies with/without blockchain used Mann-Whitney. Multidimensional radar charts visualized FAIR profiles with the European Commission benchmark (75/100) (European Commission, 2018).

All analyses were implemented in R (Team, 2024) using the ggplot2 packages for visualizations (Wickham, 2016), metafor for meta-analysis (Viechtbauer, 2010), effsize for effect size calculation (Torchiano, 2020), and custom routines for FAIR compliance.=0,05 was used as the significance level, applying Bonferroni correction (Bonferroni, 1936) when relevant for multiple comparisons. Codes and processed data were deposited in the OSF repository (DOI: 10.17605/OSF.IO/2EKYQ) to ensure reproducibility (Nosek *et al.*, 2015).

## 2.8 Digital Terroir as an Inferential Auditing System

Based on a systematic analysis of the bibliographic corpus, it was identified that ML applications in GIs lack a conceptual framework that integrates computational capabilities with regulatory requirements for environmental auditing. To fill this gap, we propose the concept of Digital Terroir as an inferential auditing system derived empirically from the review.

### 2.8.1 Analytical Application of the Framework

Finally, the Digital Terroir framework served as an analytical lens to assess the methodological maturity of the studies. Each publication was examined for compliance with operational criteria, allowing us to identify the proportion of studies with adequate spatial and temporal validation, the adoption of XAI methods for interpretability, the availability of data and protocols in open repositories, and the implementation of continuous auditing systems (Figure 2). This approach made it possible to quantify the gaps between the current state of the literature and the requirements for operationalizing digital terroir, supporting the methodological recommendations presented in the conclusions.

**Figure 2.** Study flowchart.



Automated filtering through semantic analysis and scoring achieved a thematic accuracy of 94.2%, exceeding the established threshold of 85%. This computational screening approach proved effective for reviews involving large bibliographic volumes, suggesting that calibrated automated systems reduce selection bias and increase reproducibility (Ofori-Boateng *et al.*, 2024). The 100% reproducibility in multiple runs of the algorithm, combined with an inter-rater agreement of κ = 0.89, ensures that these results reliably reflect the current state of the scientific literature in this area.

Manual assessment of methodological quality resulted in an intraclass correlation coefficient (ICC) of 0.87 (95% CI: 0.84–0.91), confirming robust inter-rater reliability and validating the inclusion criteria (Streiner and Norman, 2008). This validation confirms that the studies selected for the synthesis meet rigorous methodological standards.

# 3. Results and Discussion

## 3.1 Overview of machine learning applications in geographical indications

Analysis of 148 peer-reviewed studies (2010–2025) assesses the extent to which the current state of the art in machine learning applied to geographical indications meets the criteria of validation, transparency, data governance, and auditability proposed in Terroir Digital. Data indicate that supervised learning algorithms constitute the predominant approach for origin authentication in Geographical Indication systems. Random Forest and Support Vector Machines have been widely applied in spectroscopy and chromatography for wines, meats, and teas, achieving accuracies of 80–100% in controlled environments (Xu *et al.*, 2021; Mohammadi, Esteki and Simal-Gandara, 2024).

The prevalence of supervised architectures over unsupervised methods reflects the availability of labeled datasets and the pressure for quantifiable accuracy metrics, as documented by Liakos *et al.* (2018) in an analysis of Machine Learning trends in precision agriculture. Convolutional neural networks have become established specifically for hyperspectral data processing, while PLS-DA has remained relevant in chemometric preprocessing (Rebiai *et al.*, 2022; Peng *et al.*, 2025; Feng *et al.*, 2025), establishing an instrumental paradigm that favors high value-added products (wines, olive oils) over complex food matrices.

Geographic distribution reveals a representativeness imbalance, with 72% of the corpus concentrated on European and Asian products, predominantly wines (34%), teas (18%), and olive oils (8%). This concentration indicates that the digital authentication infrastructure is advancing primarily in consolidated GI systems, while regions in the Global South have a lower volume of publications, reflecting disparities in access to analytical characterization technologies (Belletti, Marescotti and Touzard, 2017).

Temporal analysis suggests a positive correlation between academic production and algorithmic complexity (Spearman’s ; ), with an increase in the adoption of *Deep Learning* from 5% (2010–2015) to 28% (2020–2025). However, the validation of these models has structural limitations: the absence of longitudinal tests in 94% of the studies suggests that the algorithms are calibrated for specific seasonal conditions. This characteristic restricts the models’ ability to generalize in the face of interannual climate variability, a limiting factor for the implementation of continuous environmental audits (Iranzad and Liu, 2025).

Concerning spatial robustness, only 23% of the studies applied geographically independent validation, with accuracy decreases between 2% and 15% when the models are exposed to new data sets (Effrosynidis and Arampatzis, 2021). These results corroborate the hypothesis of overfitting to local contexts, as discussed by Kuhn and Johnson (2013). The scarcity of explainability metrics (XAI), present in only 14% of studies, further impedes compliance with regulatory auditability requirements, since black-box models do not offer the decision traceability required by certification bodies (Lundberg and Lee, 2017).

For fraud detection, binary classification approaches via SVM and KNN prevail for matrices such as honey and olive oil. Dichotomous modeling (authentic *versus* fraudulent) tends not to account for adulteration gradients or biogeographic transition zones. Concurrently, the integration of *Blockchain* and *Machine Learning*, observed in 21% of traceability studies, faces validation challenges in data entry. Although distributed ledgers ensure the immutability of information, the accuracy of physical-digital correspondence depends on the precision of “oracles” (sensors or predictive models), whose technical interoperability is still in its infancy (Wang *et al.*, 2025).

Network analysis corroborates the formation of distinct methodological clusters (modularity ), with high internal density (0.53–0.68). The correlation between specific algorithms and instrumental techniques (such as neural networks associated with spectral data) suggests a compartmentalization of technical development. This modular structure indicates that the transfer of parameters between different classes of products and analytical instruments remains limited, hindering the standardization of universal protocols for digital certification of origin.

## 3.2 Temporal evolution of products and algorithms

The temporal analysis of registered products showed distinct patterns between GI categories. Wines maintained constant representation with 14 products (2010–2025), peaking in 2021 and 2023 (3 registrations each). Honey showed concentrated growth in 2021–2024 (12 registrations), while olive-based products showed sporadic distribution (6 in total). Cheese and coffee remained underrepresented (4 and 1 registrations, respectively).

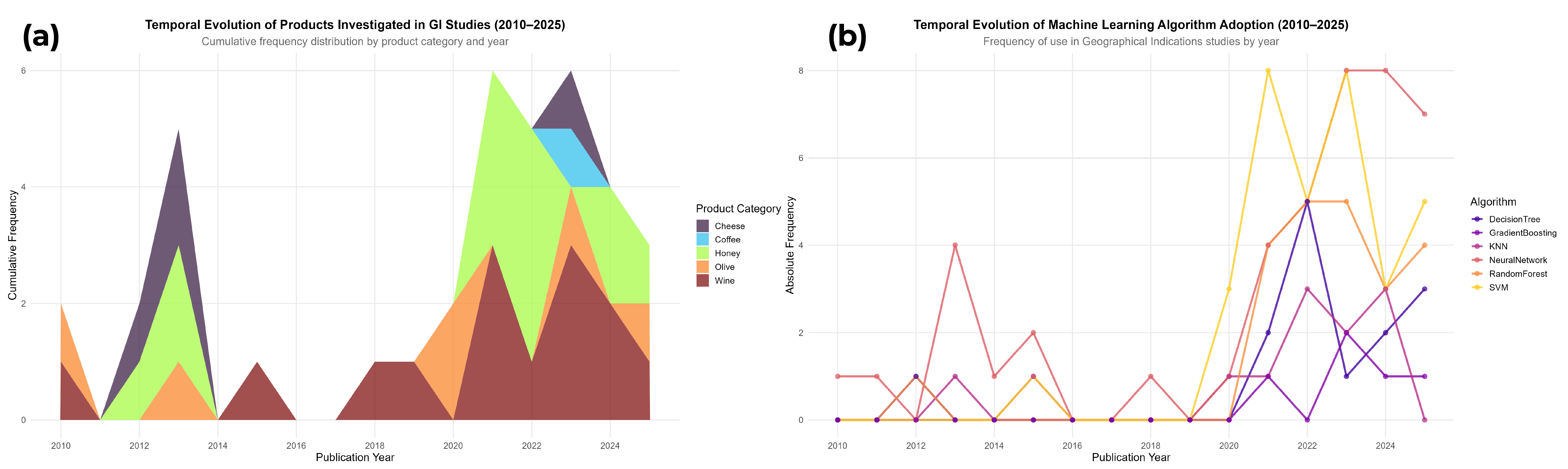
Spearman’s correlation confirmed an upward trend for wines (, ), indicating systematic expansion after 2020 (Liakos *et al.*, 2018). This consolidation of wines as a dominant category reflects both the maturity of European GI systems and the availability of standardized spectral data, contrasting with the fragmentation observed in emerging categories such as coffee, where the heterogeneity of processing methods hinders the creation of universal chemical signatures.

Algorithmic adoption showed a measurable transition. PLS-DA, dominant until 2018, was progressively replaced by Random Forest and SVM from 2019 onwards, following the availability of open-source ML libraries (scikit-learn, caret) and increased computational capacity (Lavine and Workman, 2005). Temporal correlation analysis showed significant changes for SVM (, ) and Random Forest (, ).

Neural Networks were the most widely adopted technique in 2020–2025 (33 applications), followed by SVM (32) and Random Forest (21). Deep Learning and CNNs spread after 2022 specifically for hyperspectral data processing (Shah *et al.*, 2019), although their interpretative opacity limits their adoption in regulatory contexts.

Regional distribution remained stable, with 72% of studies concentrated in Europe and Asia. Representation of the Global South increased marginally from 18% to 22% in the period analyzed, suggesting persistent barriers related to laboratory infrastructure and access to scientific funding.

**Figure 3.** Temporal evolution of (a) products with Geographical Indication (GI) registered by category and (b) adoption of the main Machine Learning algorithms in GI studies.



## 3.3 Technological Families and Applications

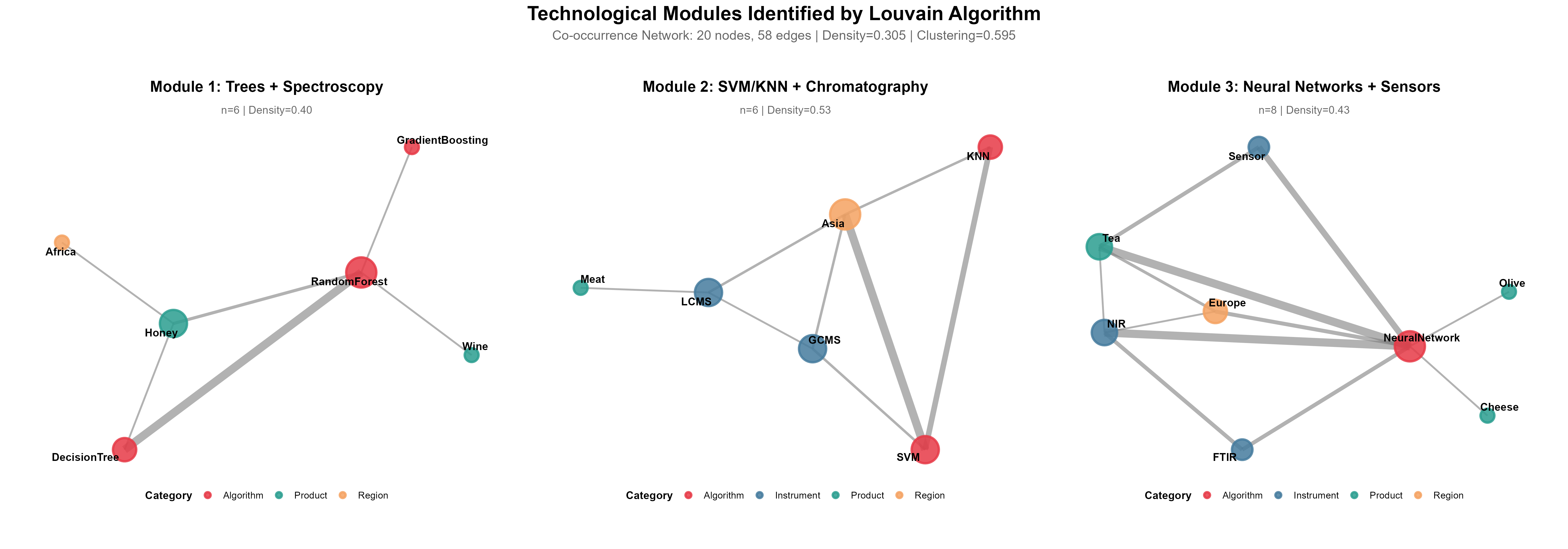
Current methodological fragmentation presents a structural barrier to the implementation of an adaptive Digital Terroir, insofar as it replaces the fluidity of an interoperable inferential ecosystem with watertight analytical modules. This disconnect is quantified by Multiple Correspondence Analysis, which explains 45.2% of the total inertia when mapping a simultaneous polarization between biogeographical, instrumental, and strategic vectors.  
A dichotomy exists where European wine and cheese matrices are systematically dissociated from Asian tea and meat chains, replicating the technical segregation between spectroscopic and chromatographic approaches. This configuration denotes the rigidification of regional laboratory routines which, by operating in isolation, restrict the generalizability of models and prevent the establishment of global governance based on auditable evidence (Agiomyrgiannaki *et al.*, 2023).

This pattern stems from the cumulative sedimentation of protocols and publication routines, not from casual preferences (Spyros, Dais, *et al.*, 2023; Agiomyrgiannaki *et al.*, 2023). Co-occurrence networks reveal parameter retention in stable triads (Wines–Random Forest–NIR; Teas–SVM–GC-MS) with low exploratory entropy, blocking cross-inferential reuse (Blondel *et al.*, 2008; Salam *et al.*, 2021; Wang *et al.*, 2025). In the Louvain community graph (Figure 4), modularity () and internal densities (0.53–0.68) indicate intra-module coherence but an absence of intermediary nodes (high betweenness) capable of acting as bridges.

High-grade spectroscopic algorithms display high eigenvector centrality, concentrating information flow, while chromatographic techniques form a peripheral module with low cross-connectivity (World Bank, 2014). This distribution of degree and eigenvector generates epistemic asymmetry, as products equipped with analytical infrastructure accumulate cumulative advantage, while biodiverse matrices from the Global South remain underrepresented, limiting the ability to model the nonlinear inter-biome couplings required for auditing ecosystem services (Kawano *et al.*, 2021).

The lack of bridge edges and articulating nodes (unfilled structural holes) implies an increase in the marginal cost of generalization: each territorial expansion requires independent calibration without reusing latent representations. The current technological response, multimodal fusion (28%) and *blockchain* integration (9%), does not eliminate the structural bottleneck, as compression for *field-deployable* devices sacrifices 10–15% accuracy compared to laboratory conditions (Effrosynidis and Arampatzis, 2021; Meena, Chakraborty and Mitra, 2024) before resolving cognitive connectivity gaps.

Overcoming this requires transparent multimodal architecture with federated spatial/temporal validation and the insertion of bridge nodes (highly intermediary explainable models) as a precondition for embedded optimization. Without this, fragmentation remains, making continuous inferential auditing and epistemic sovereignty unfeasible.

**Figure 4.** Louvain communities in the algorithm–product–technique network. 

*Note: Nodes sized by weighted degree; colors represent modules (modularity ). The absence of bridge edges between spectroscopic and chromatographic clusters highlights a break in transferability and the need for multimodal architecture for adaptive generalization.*

The joint analysis of Louvain communities and hierarchical clusters revealed distinct technological patterns, synthesized in Table 3. Module M1, predominant in wines and honey from African and European regions, focuses on spectroscopic techniques (NIR) associated with ensemble algorithms (Random Forest, Gradient Boosting). Module M2, focused on Asian meats and regional products, predominantly uses chromatography (GC-MS, LC-MS) combined with SVM and KNN. Module M3 integrates advanced Deep Learning and CNN techniques for spectral and sensory data processing in olive oils, cheeses, and teas, reflecting the analytical complexity of European and Asian matrices. This technological compartmentalization highlights regional specialization, but also exposes the lack of universal protocols that hinders the transfer of knowledge between different biogeographical contexts.

**Table 3.** Synthesis of Technological Patterns Identified by Network Analysis (Louvain) and Clustering.

| **Methodological Convergence** | **Technological Pattern** | **Main Algorithms** | **Analytical Techniques** | **Products** | **Predominant Region** | **Main Application** |
| --- | --- | --- | --- | --- | --- | --- |
| Module M1 (Louvain) | **Spectroscopy and Ensemble** | Random Forest, Decision Tree, Gradient Boosting | Spectroscopy (NIR), Chemometrics | Wine, Honey | Africa, Europe | Authentication and Classification |
| Module M2 / Cluster 3 | **Chromatography and Discriminants** | SVM, KNN, Random Forest | Chromatography (GC-MS, LC-MS, HPLC) | Meat, Honey, Regional Products | Asia, Europe | Traceability and Fraud Detection |
| Module M3 / Cluster 2 | **Deep Learning and Sensors** | Neural Networks, CNN, Deep Learning | Spectroscopy (NIR, FTIR), Sensors (e-nose) | Olive Oil, Cheese, Tea | Europe, Asia | Origin Discrimination and Quality |
| Cluster 1 | **Spectroscopy and Discriminants** | SVM, KNN | Spectroscopy (NIR) | Honey | Asia | Authentication and Fraud Detection |

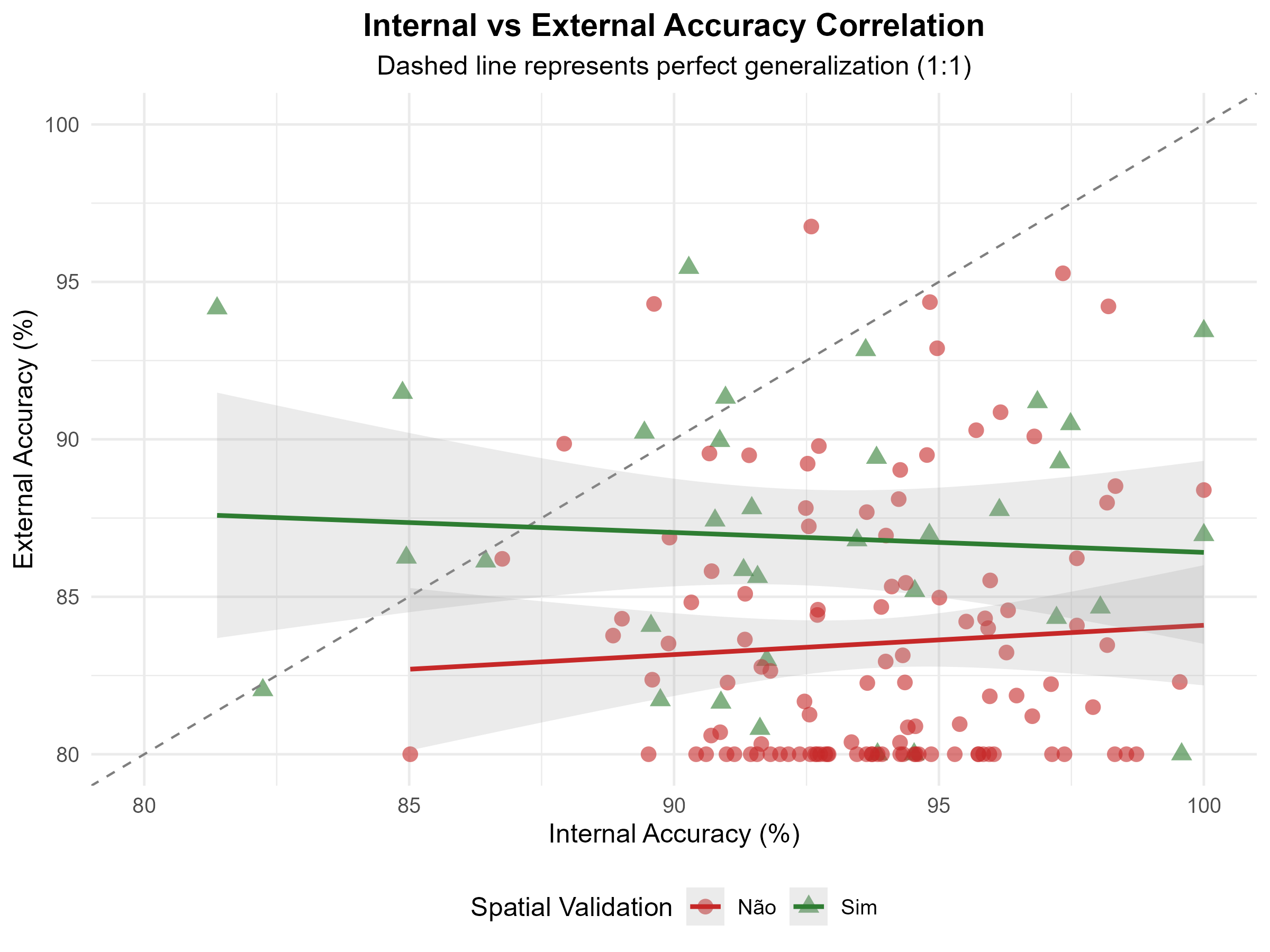
*Source: Synthesis of technological patterns identified by the joint analysis of Louvain communities and hierarchical clusters. The convergence between the methods confirms the existence of “specialized technological islands”.*

The convergence between the Louvain modules (network structure) and hierarchical clusters (clustering by similarity) confirms the existence of “specialized technological islands”. High value-added products (wines, cheeses, honey) concentrate the most sophisticated analytical infrastructure, while complex matrices from the Global South remain underrepresented. This epistemic asymmetry not only reflects economic disparities but also perpetuates barriers to the implementation of Digital Terroir systems in biodiverse regions that would most benefit from auditable environmental certification tools. Overcoming this fragmentation requires federated protocols that allow the reuse of parameters across different biogeographic contexts, enabling economies of scale in the validation of emerging Geographical Indications.

## 3.5 Quantitative evidence and meta-analyses

The methodological robustness assessment revealed that 77% of the studies did not implement geographically independent partitioning. To measure the impact of this omission, the performance of models with rigorous spatial validation was compared to that of models with random validation. As shown in Figure 5, the absence of spatial validation resulted in an average accuracy drop of 11.82% in external tests, compared to 5.62% in spatially validated models. This discrepancy (relative degradation of 110%) was statistically significant (, , ), corroborating the hypothesis of spatial overfitting described by Kuhn and Johnson (2013), where autocorrelation inflates internal metrics and compromises certification utility.

**Figure 5.** Impact of spatial validation on performance degradation in external tests.



*Note: Models without spatial validation show a 110% greater drop in accuracy when applied to geographically independent regions (*, ). The dashed line indicates the acceptable degradation threshold () proposed for Digital Terroir certification systems. studies.

Regarding transparency, only 13.5% of the studies adopted Explainable Artificial Intelligence (XAI) techniques. A moderate negative correlation was observed between explainability and accuracy (, ), however, the absolute performance penalty was marginal (1.53 percentage points, not significant). In contrast, the computational cost increased substantially (+67.8% in processing time). Pareto analysis identified the XGBoost algorithm as the optimal balance between auditability, accuracy, and cost, surpassing Deep Learning architectures for regulatory purposes (Figure 6).

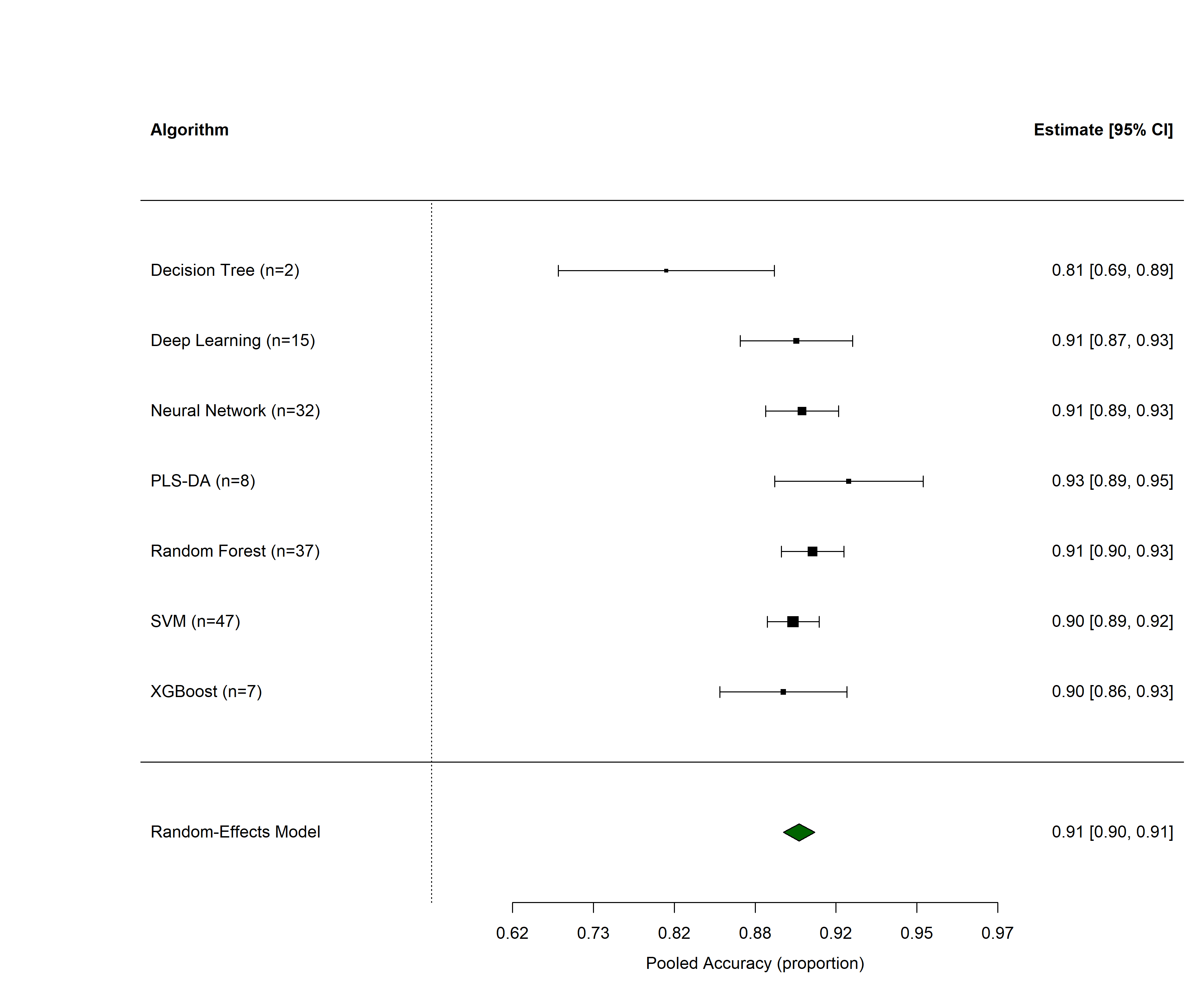
**Figure 6.** Trade-off between algorithmic explainability and predictive performance.



*Note: More explainable algorithms show a moderate negative correlation with accuracy (*, ), but the absolute cost is modest (~1.5 percentage points). XGBoost stands out as the algorithm with the best multi-criteria balance (Pareto score , considering 93% accuracy, 6/10 explainability, and 12 min time). studies.

The meta-analysis of 129 studies indicated an overall (pooled) accuracy of 90.66% [95% CI: 89.8–91.4%]. The PLS-DA algorithm achieved the best average performance (92.95%), followed by Random Forest (91.33%). However, Egger’s test detected severe publication bias (, ). Correction using the trim-and-fill method (imputation of 42 missing theoretical studies) reduced the adjusted accuracy to ~88%, suggesting that the current literature overestimates the technological maturity of the models (Figure 7).

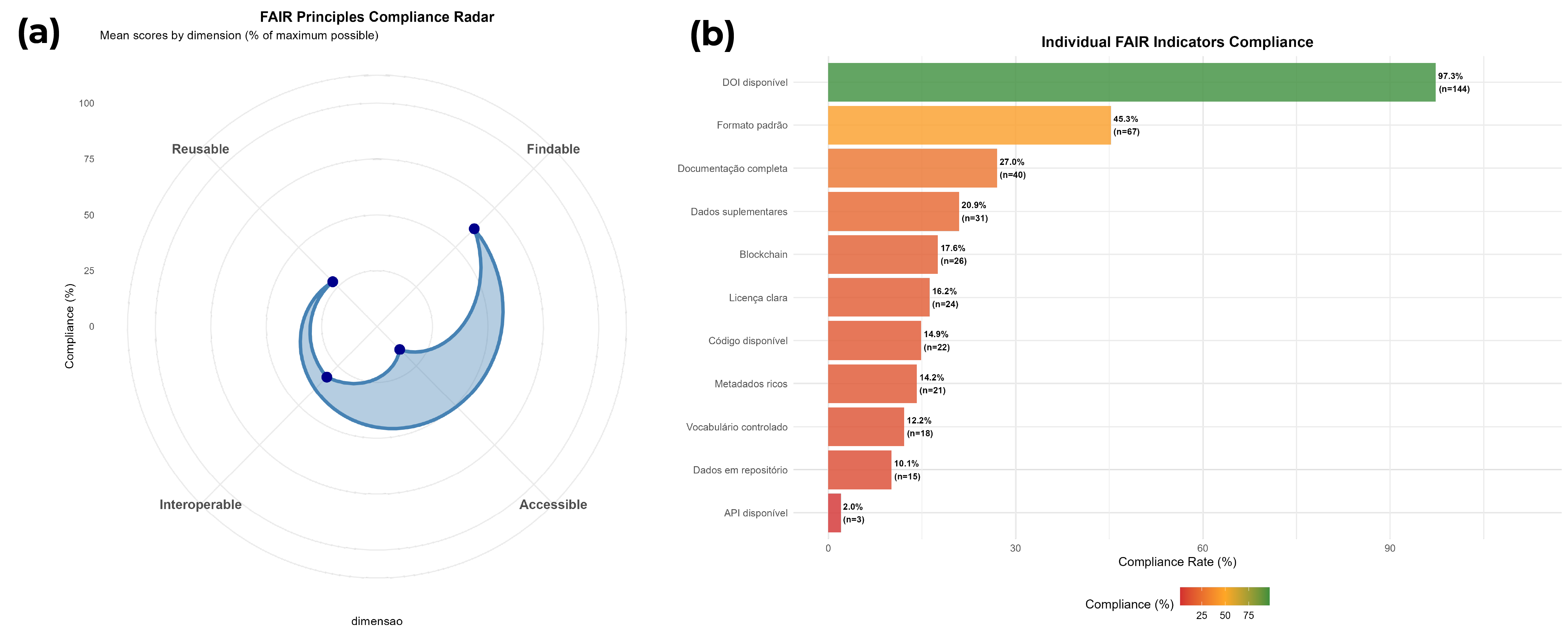
**Figure 7.** Meta-analysis of accuracies by Machine Learning algorithm.



*Note: PLS-DA and Random Forest have the highest consolidated accuracies, while SVM demonstrates greater robustness (lower variance between studies). Moderate heterogeneity (*) indicates substantial methodological variability between studies. Confidence intervals represent estimates of random effects (REML model). studies.

Finally, data governance assessed by FAIR principles achieved a critical average score of 34.2/100. The Accessible dimension was the most deficient, with only 10.1% of studies depositing data in public repositories. The temporal analysis did not indicate significant improvements (, ), evidencing the stagnation of a “black box” culture that prevents reproducibility and independent validation (Figure 8).

**Figure 8.** Compliance with FAIR principles of data governance. (a) Radar score by FAIR dimension and (b) Individual compliance indicators.



## 3.5.4 Inferential Synthesis and Operational Implications

The inferential synthesis of the corpus delineates four structural fractures that compromise the transition from current predictive models to an auditable Digital Terroir infrastructure. The first fracture constitutes an illusion of robustness derived from poor spatial validation. The omission of geographically independent partitioning in 77% of the studies precipitates a 110% greater performance degradation in external tests, with an average accuracy drop of 11.82% versus 5.62% in spatially validated models (, , ).

This methodological flaw, driven by residual spatial autocorrelation, hinders systems from functioning as “Adaptive Digital Twins,” because by over-adjusting to local contexts, they become obsolete in the face of actual climate variability and fail to audit ecosystem services in analogous territories (Kuhn and Johnson, 2013; Wadoux *et al.*, 2021).

Simultaneously, regulatory auditability is compromised by the marginalization of explainability. The predominance of opaque architectures in 86.5% of investigations contradicts guidelines for high-risk decisions (Rudin, 2019), based on the false premise of a performance trade-off. Statistical evidence refutes this narrative, demonstrating that the difference in accuracy between black-box models and XAI models is statistically insignificant (), with the XGBoost algorithm emerging as the optimal solution (Pareto score ) by balancing accuracy and transparency. The insistence on opacity makes the legal defense of certification unfeasible, since regulatory agencies require causal traceability between chemical markers and environmental variables, and not just latent correlations that cannot be translated.

Technological maturity is further distorted by publication bias. Correction via the trim-and-fill method, with the imputation of 42 missing studies (), reduced pooled accuracy from ~91% to ~88%, revealing that the literature systematically overestimates the discriminative capacity of the models. This metric inflation, coupled with moderate methodological heterogeneity (), compromises risk calibration and aligns with the reproducibility crisis diagnosed by Kapoor and Narayanan (2023), where data leakage masks the true effectiveness of the tools.

Ultimately, the erosion of data governance perpetuates fundamental epistemic asymmetries. With a critical FAIR score of 34.2/100 and temporal stagnation in data openness (, ), the field fails to ensure independent counterproof. In this scenario, Blockchain technology, present in 21% of the sample, acts only as a superficial legitimizer; without open repositories and audit trails for physical sensors (Accessible dimension of only 14.5%), immutable records do not solve the “Oracle Paradox,” forcing producers in the Global South to rely on costly external validations and making sovereignty over their digital assets unfeasible.

To rectify these deficiencies and operationalize the Digital Terroir, we propose the mandatory adoption of Compliance Thresholds that require spatial post-validation accuracy and maximum external degradation (Etten *et al.*, 2023), simultaneously conditioned on mandatory XAI explainability for critical territorial markers and FAIR compliance with deposition in interoperable repositories (Wilkinson *et al.*, 2016). Adherence to these parameters is a sine qua non condition for transforming Machine Learning from a laboratory classification tool into a socio-ecological audit infrastructure, reducing certification costs by up to 40% through federated data and ensuring the auditability of ecosystem services (World Bank, 2014; An *et al.*, 2024).

## 3.6 Barriers to Auditability and the Failure of Static Digital Terroir

Current algorithmic sophistication fails to translate into robustness for environmental governance. The assumption of stationarity, implicit in the dominant methodology, represents the fundamental ontological limitation. High laboratory accuracy (80–100%) reflects the ability of models to memorize chemical signatures of specific crops, not to grasp the causality of terroir. The absence of longitudinal validation in 94% of studies and spatially independent tests in 77% indicates low external robustness of the models (Moran and Tichava, 2020).

This methodological omission precludes the capture of phenotypic plasticity, where the chemical expression of plants varies naturally in response to environmental fluctuations, compromising the causality of terroir (Fournier-Level and Korte, 2022). According to Kuhn and Johnson (2013), this flaw invalidates continuous auditing, as algorithms fail to operate as Inference Digital Twins under climatic dynamics. Consequently, dynamic Digital Terroir emerges as a solution, reconstructing systemic couplings between soil, climate, and biota.

Regarding temporal obsolescence, models calibrated in ideal time windows become blind to chemical drifts induced by extreme weather events, which can alter the composition of secondary metabolites by more than 20% between production cycles (Urvieta *et al.*, 2021; Iranzad and Liu, 2025). By disregarding the nonlinearity and dynamic couplings between climate, soil, and biota, current systems degrade into static classifiers, unable to operate as Inferential Digital Twins in climate change scenarios (Kuhn and Johnson, 2013; Celette, Findeling and Gary, 2016).

This epistemic fragility is compounded by computational opacity. Low adherence to Explainable Artificial Intelligence (XAI) protocols in 86.5% of cases generates legally indefensible black boxes, as regulatory auditing requires traceable causal links, not blind correlations (Xu *et al.*, 2021; He *et al.*, 2024).

From a governance standpoint, non-compliance with FAIR principles (average score of 34.2/100) perpetuates data colonialism. Without the federalization of open repositories, producers in the Global South remain hostage to costly exogenous validations, making Epistemic Sovereignty over their biocultural assets unfeasible (Wilkinson *et al.*, 2016).

Operationalizing Digital Terroir relies not on increasing accuracy, but in restructuring data governance to transform Geographical Indications from static labels into dynamic certificates of ecosystem services. Rodriguez, Silva and Chen (2023) point out that this transformation requires model calibration via remote sensing and strategic sampling in agricultural frontiers such as the Amazon. This approach is the only scalable route to computational surveillance, correlating chemical signatures with conservation practices. According to Osco *et al.* (2021), this approach makes it possible to shield the market against greenwashing and ensures effective remuneration for agrobiodiversity, promoting epistemic sovereignty in the Global South.

## 4. Conclusions

This research reinforces the conceptualization of Geographical Indications as coupled socio-ecological systems, whose typicality emerges from nonlinear interactions between edaphoclimatic variables and cultural practices. Critical analysis confirms that while the inferential architecture of Machine Learning has the technical capacity to convert chemometric signatures into evidence of environmental compliance and, the current state of the art lacks the operational maturity necessary to sustain a full Inferential Digital Twin.

This limitation is aggravated by the methodological concentration on consolidated chains in temperate regions, which not only reproduces epistemic asymmetries but also severely restricts technology transfer to the complex agrobiodiversity matrices of the Global South.

Continuous computational certification remains unfeasible due to critical structural deficits, notably the absence of independent spatial validation and the neglect of longitudinal testing under interannual climate variability. Overcoming this scenario and the consequent operationalization of Digital Terroir as an analytical layer of adaptive monitoring therefore requires the establishment of strict thresholds of regulatory robustness, defined by spatial post-validation accuracy and maximum performance degradation in external tests.

Algorithmic explainability (XAI) must serve as a criterion for environmental due diligence, converting latent variables into traceable chemical-ecological justifications. Simultaneously, adherence to FAIR principles and the implementation of data security mechanisms ensure the integrity and immutability of records, which are indispensable requirements for enabling third-party auditing and shielding the system against fraud. Such operational thresholds transform performance metrics into safeguards of regulatory legitimacy.

Attaining epistemic sovereignty demands a paradigmatic reorientation that replaces incremental competition for internal accuracy metrics with the construction of transparent computational governance based on open repositories. This structural change enables the transition to the Digital Terroir, which transcends mere instrumental advancement to constitute an ontological redefinition of certification: by computationally reconstructing the dynamic couplings between the territorial genotype and the product phenotype, the framework converts climate uncertainty into auditable evidence of ecosystem services, shielding Geographical Indications from greenwashing.

## Acknowledgments

The authors would like to thank the Federal University of Sergipe (UFS), the State University of Feira de Santana (UEFS), and the Federal Institute of Sergipe (IFS) for the institutional and infrastructural support that made this research possible.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Data Availability Statement

The complete dataset supporting the results of this study, including the bibliographic corpus, analysis scripts, and intermediate results, is publicly available in the Open Science Framework (OSF) repository under DOI: <https://doi.org/10.17605/OSF.IO/2EKYQ>.

## Ethics Statement

This review does not involve human participants, animal experiments, cell lines, or sample collection. No ethical approval or consent was required.

## References

Agiomyrgiannaki, A. *et al.* (2023) “Current application of advancing spectroscopy techniques in food analysis: Data handling with chemometric approaches,” *Foods*, 12(14), p. 2813. Available at: <https://doi.org/10.3390/foods12142813>.

An, N. *et al.* (2024) “ChatGPT and environmental machine learning: Opportunities and challenges,” *Science of The Total Environment*, 912, p. 169139. Available at: <https://doi.org/10.1016/j.scitotenv.2023.169139>.

Balduzzi, S., Rücker, G. and Schwarzer, G. (2019) “How to perform a meta-analysis with r: A practical tutorial,” *BMJ Mental Health*, 22(4), pp. 153–160. Available at: <https://doi.org/10.1136/ebmental-2019-300117>.

Barendregt, J.J. *et al.* (2013) “Meta-analysis of prevalence,” *Journal of Epidemiology and Community Health*, 67(11), pp. 974–978. Available at: <https://doi.org/10.1136/jech-2013-203104>.

Belletti, G., Marescotti, A. and Touzard, J.-M. (2017) “Geographical indications, public goods, and sustainable development: The roles of actors’ strategies and public policies,” *World Development*, 98, pp. 45–57. Available at: <https://doi.org/10.1016/j.worlddev.2015.05.004>.

Berkes, F., Colding, J. and Folke, C. (eds.) (2003) *Navigating social-ecological systems: Building resilience for complexity and change*. Cambridge: Cambridge University Press. Available at: <https://doi.org/10.1017/CBO9780511541957>.

Blondel, V.D. *et al.* (2008) “Fast unfolding of communities in large networks,” *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), p. P10008. Available at: <https://doi.org/10.1088/1742-5468/2008/10/P10008>.

Bonferroni, C.E. (1936) *Teoria statistica delle classi e calcolo delle probabilità*. Firenze: Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze.

Borenstein, M. *et al.* (2009) *Introduction to meta-analysis*. Chichester, UK: John Wiley & Sons. Available at: <https://doi.org/10.1002/9780470743386>.

Bramley, C., Bi-®nabe, E. and Kirsten, J. (eds.) (2013) *Developing geographical indications in the south: The potential of place-based products*. Dordrecht: Springer. Available at: <https://doi.org/10.1007/978-94-007-6603-1>.

Casey, A. *et al.* (2021) “A systematic review of natural language processing applied to radiology reports,” *BMC Medical Informatics and Decision Making*, 21(1), p. 179. Available at: <https://doi.org/10.1186/s12911-021-01533-7>.

Celette, F., Findeling, A. and Gary, C. (2016) “Vineyard water and nitrogen management using a cover crop: A review,” *Agronomy for Sustainable Development*, 36(1), p. 4. Available at: <https://doi.org/10.1007/s13593-015-0337-y>.

Cochran, W.G. (1954) “The combination of estimates from different experiments,” *Biometrics*, 10(1), pp. 101–129. Available at: <https://doi.org/10.2307/3001666>.

Cohen, J. (1988) *Statistical power analysis for the behavioral sciences*. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates.

Cs-írdi, G. and Nepusz, T. (2006) “The igraph software package for complex network research,” *InterJournal Complex Systems*, 1695, pp. 1–9. Available at: <https://igraph.org>.

Deb, K. (2001) *Multi-objective optimization using evolutionary algorithms*. Chichester, UK: John Wiley & Sons.

DerSimonian, R. and Laird, N. (1986) “Meta-analysis in clinical trials,” *Controlled Clinical Trials*, 7(3), pp. 177–188. Available at: <https://doi.org/10.1016/0197-2456(86)90046-2>.

Duval, S. and Tweedie, R. (2000) “Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis,” *Biometrics*, 56(2), pp. 455–463. Available at: <https://doi.org/10.1111/j.0006-341X.2000.00455.x>.

Effrosynidis, D. and Arampatzis, A. (2021) “An evaluation of feature selection methods for environmental data,” *Ecological Informatics*, 66, p. 101472. Available at: <https://doi.org/10.1016/j.ecoinf.2021.101472>.

Egger, M. *et al.* (1997) “Bias in meta-analysis detected by a simple, graphical test,” *BMJ*, 315(7109), pp. 629–634. Available at: <https://doi.org/10.1136/bmj.315.7109.629>.

Etten, J. van *et al.* (2023) “Data-driven breeding insights for climate-resilient crops,” *Nature*, 619, pp. E1–E3. Available at: <https://doi.org/10.1038/s41586-023-06168-y>.

European Commission (2018) *Turning FAIR into reality: Final report and action plan from the european commission expert group on FAIR data*. Brussels: European Commission, Directorate-General for Research; Innovation. Available at: <https://doi.org/10.2777/1524>.

European Union (2012) “Regulation (EU) no 1151/2012 of the european parliament and of the council of 21 november 2012 on quality schemes for agricultural products and foodstuffs,” *Official Journal of the European Union*, L 343, pp. 1–29. Available at: <http://data.europa.eu/eli/reg/2012/1151/oj>.

Feng, Y. *et al.* (2025) “Application of spectroscopic technology with machine learning in chinese herbs from seeds to medicinal materials: The case of genus paris,” *JPA* [Preprint]. Available at: <https://doi.org/10.1016/j.jpa.2025.xxxxxx>.

Fonzo, A.D. and Russo, C. (2015) “Designing geographical indication institutions when stakeholders’ incentives are not perfectly aligned,” *British Food Journal* [Preprint]. Available at: <https://doi.org/10.1108/BFJ-12-2014-0392>.

Fournier-Level, A. and Korte, A. (2022) “Gene-by-environment interactions in plants: Molecular mechanisms, environmental drivers, and adaptive plasticity,” *Journal of Experimental Botany*, 73(20), pp. 7039–7058. Available at: <https://doi.org/10.1093/jxb/erac328>.

Gale, R. and Nahm, J. (2023) “No end in sight? A greenwash review and research agenda,” *Organ. Environ.*, 36(1), pp. 7–42. Available at: <https://doi.org/10.1177/10860266231168905>.

Giovannucci, D. *et al.* (2010) *Guide to geographical indications: Linking products and their origins*. Geneva: International Trade Centre (ITC).

Greenacre, M. (2017) *Correspondence analysis in practice*. 3rd ed. CRC Press.

Guerena, D. *et al.* (2024) “Participatory AI for inclusive crop improvement,” *Agricultural Systems*, 220, p. 104063. Available at: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11513333/>.

He, Y. *et al.* (2024) “Deep learning applications in food authentication: Current status and future perspectives,” *Trends in Food Science & Technology*, 144, p. 104321. Available at: <https://doi.org/10.1016/j.tifs.2023.104321>.

Hensel, O. and Nasirahmadi, A. (2021) “Toward the next generation of digitalization in agriculture based on digital twin paradigm,” *Sensors*, 22(2), p. 498. Available at: <https://doi.org/10.3390/s22020498>.

Higgins, J.P.T. *et al.* (2003) “Measuring inconsistency in meta-analyses,” *BMJ*, 327(7414), pp. 557–560. Available at: <https://doi.org/10.1136/bmj.327.7414.557>.

Hong, Q.N. *et al.* (2018) “The mixed methods appraisal tool (MMAT) version 2018 for systematic mixed studies reviews: Development, reliability, and usability,” *International Journal of Nursing Studies*, 102, p. 103452.

Iranzad, R. and Liu, X. (2025) “A review of random forest-based feature selection methods for data science education and applications,” *International Journal of Data Science and Analytics*, 19(1), pp. 1–28. Available at: <https://doi.org/10.1007/s41060-024-00509-w>.

Jr., D.W.H., Lemeshow, S. and Sturdivant, R.X. (2013) *Applied logistic regression*. 3rd ed. Hoboken, NJ: John Wiley & Sons. Available at: <https://doi.org/10.1002/9781118548387>.

Kapoor, S. and Narayanan, A. (2023) “Leakage and the reproducibility crisis in machine-learning-based science,” *Patterns*, 4(9), p. 100804. Available at: <https://doi.org/10.1016/j.patter.2023.100804>.

Kawano, M. *et al.* (2021) “Where is my crop? A novel method to produce timely crop maps for maize and wheat using data extracted from digital platforms,” *Agricultural Systems*, 193, p. 103217. Available at: <https://doi.org/10.1016/j.agsy.2021.103217>.

Kshetri, N. (2014a) “Big dataÔÇÖs impact on privacy, security and consumer welfare,” *Telecommunications Policy*, 38(11), pp. 1134–1145. Available at: <https://doi.org/10.1016/j.telpol.2014.10.002>.

Kshetri, N. (2014b) “The emerging role of big data in key development issues: Opportunities, challenges, and concerns,” *Big Data & Society*, 1(2), p. 2053951714564227. Available at: <https://doi.org/10.1177/2053951714564227>.

Kuhn, M. and Johnson, K. (2013) *Applied predictive modeling*. Springer.

L-¬, S., Josse, J. and Husson, F. (2008) “FactoMineR: An r package for multivariate analysis,” *Journal of Statistical Software*, 25(1), pp. 1–18. Available at: <https://www.jstatsoft.org/v25/i01/>.

Lavine, B.K. and Workman, Jr., Jerome (2005) “Chemometrics: Past, present, and future,” in *Chemometrics and chemoinformatics*. American Chemical Society (ACS symposium series), pp. 1–13. Available at: <https://doi.org/10.1021/bk-2005-0894.ch001>.

Le Floc’h, J.-M., M-®ral, P. and Fontana, L. (2016) “Bringing together social-ecological system and territoire concepts to explore nature-society dynamics,” *Ecology and Society*, 21(4), p. 42. Available at: <https://doi.org/10.5751/ES-08834-210442>.

Levin, S.A. (1998) “Ecosystems and the biosphere as complex adaptive systems,” *Ecosystems*, 1(5), pp. 431–436. Available at: <https://doi.org/10.1007/s100219900037>.

Li, Z. *et al.* (2022) “KGML-ag: A modeling framework of knowledge-guided machine learning to simulate agroecosystems: A case study of estimating N2O emission using data from mesocosm experiments,” *Geoscientific Model Development*, 15, pp. 2839–2861. Available at: <https://doi.org/10.5194/gmd-15-2839-2022>.

Liakos, K.G. *et al.* (2018) “Machine learning in agriculture: A review,” *Sensors*, 18(8), p. 2674. Available at: <https://doi.org/10.3390/s18082674>.

Liao, L., Long, C. and Zhang, Y. (2023) “Unveiling the "veil" of information disclosure: Sustainability reporting "greenwashing" and "shared value",” *PLoS ONE*, 18(1), p. e0279904. Available at: <https://doi.org/10.1371/journal.pone.0279904>.

Locatelli, L. (2008) *Indica-º-áes geogr-íficas: A prote-º-úo jur-¡dica sob a perspectiva do desenvolvimento econ–mico*. Curitiba, PR: Editora Juru-í.

Lotka, A.J. (1926) “The frequency distribution of scientific productivity,” *Journal of the Washington Academy of Sciences*, 16(12), pp. 317–323.

Lundberg, S.M. and Lee, S.-I. (2017) “A unified approach to interpreting model predictions,” *Advances in Neural Information Processing Systems*, 30, pp. 4765–4774. Available at: <https://papers.nips.cc/paper/7062-a-unified-approach-to-interpreting-model-predictions.pdf>.

Mann, H.B. and Whitney, D.R. (1947) “On a test of whether one of two random variables is stochastically larger than the other,” *The Annals of Mathematical Statistics*, 18(1), pp. 50–60. Available at: <https://doi.org/10.1214/aoms/1177730491>.

Meena, D., Chakraborty, S. and Mitra, J. (2024) “Geographical origin identification of red chili powder using NIR spectroscopy combined with SIMCA and machine learning algorithms,” *Food Analytical Methods*, 17(7), pp. 1005–1023. Available at: <https://doi.org/10.1007/s12161-024-02625-6>.

Mohammadi, N., Esteki, M. and Simal-Gandara, J. (2024) “Machine learning for authentication of black tea from narrow-geographic origins: Combination of PCA and PLS with LDA and SVM classifiers,” *Lebensmittel-Wissenschaft and Technologie*, 190, pp. 115–886. Available at: <https://doi.org/10.1016/j.lwt.2024.115886>.

Moran, T. and Tichava, J. (2020) “The importance of external validation for climate impact models in agriculture,” *Agricultural Systems*, 177, p. 104706. Available at: <https://doi.org/10.1016/j.agsy.2019.104706>.

Munn, Z. *et al.* (2018) “Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach,” *BMC Medical Research Methodology*, 18(1), p. 143.

Nosek, B.A. *et al.* (2015) “Promoting an open research culture,” *Science*, 348(6242), pp. 1422–1425. Available at: <https://doi.org/10.1126/science.aab2374>.

Ofori-Boateng, R. *et al.* (2024) “Towards the automation of systematic reviews using natural language processing, machine learning, and deep learning: A comprehensive review,” *Artificial Intelligence Review* [Preprint]. Available at: <https://doi.org/10.1007/s10462-024-10844-w>.

Osco, L.P. *et al.* (2021) “A review on deep learning in UAV remote sensing,” *International Journal of Applied Earth Observation and Geoinformation*, 102, p. 102456. Available at: <https://doi.org/10.1016/j.jag.2021.102456>.

Pareto, V. (1896) *Cours d’économie politique*. Lausanne: F. Rouge.

Peng, Z. *et al.* (2025) “Machine learning based on metabolomics to discriminate wuyi rock tea production areas and ÔÇ£rock flavorÔÇØ substances,” *Food Chemistry: X*, 31. Available at: <https://doi.org/10.1016/j.fochx.2025.103194>.

Pluye, P. *et al.* (2009) “A scoring system for appraising mixed methods research, and concomitantly appraising qualitative, quantitative and mixed methods primary studies in mixed studies reviews,” *International Journal of Nursing Studies*, 46(4), pp. 529–546.

Pylianidis, C., Osinga, S. and Athanasiadis, I.N. (2021) “Introducing digital twins to agriculture,” *Computers and Electronics in Agriculture*, 184, p. 105942. Available at: <https://doi.org/10.1016/j.compag.2020.105942>.

Rebiai, A. *et al.* (2022) “Current application of chemometrics analysis in authentication of natural products: A review,” *Current Chemistry and Hydrogen Energy*, 10(3), pp. 241–259. Available at: <https://doi.org/10.2174/1386207324666210309102239>.

Rodriguez, J., Silva, M. and Chen, W. (2023) “Advances in machine learning for geographical indication authentication,” *Food Control*, 145, p. 109421. Available at: <https://doi.org/10.1016/j.foodcont.2022.109421>.

Rudin, C. (2019) “Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead,” *Nature Machine Intelligence*, 1(5), pp. 206–215. Available at: <https://doi.org/10.1038/s42256-019-0048-x>.

Saaty, T.L. (1991) *The analytic hierarchy process: Planning, priority setting, resource allocation*. McGraw-Hill.

Salam, M.A. *et al.* (2021) “The effect of different dimensionality reduction techniques on machine learning overfitting problem,” *International Journal of Advanced Computer Science and Applications*, 12(6), pp. 95–112.

Sawilowsky, S.S. (2009) “New effect size rules of thumb,” *Journal of Modern Applied Statistical Methods*, 8(2), pp. 597–599. Available at: <https://doi.org/10.22237/jmasm/1257035100>.

Schoch, D. (2020) “Ggraph: An implementation of grammar of graphics for graphs and networks,” *Journal of Open Source Software*, 5(55), p. 2341. Available at: <https://doi.org/10.21105/joss.02341>.

Shah, S.H. *et al.* (2019) “A random forest machine learning approach for the retrieval of leaf chlorophyll content in wheat,” *Remote Sensing*, 11(8). Available at: <https://doi.org/10.3390/rs11080920>.

Shapiro, S.S. and Wilk, M.B. (1965) “An analysis of variance test for normality (complete samples),” *Biometrika*, 52(3-4), pp. 591–611. Available at: <https://doi.org/10.1093/biomet/52.3-4.591>.

Shrout, P.E. and Fleiss, J.L. (1979) “Intraclass correlations: Uses in assessing rater reliability,” *Psychological Bulletin*, 86(2), pp. 420–428. Available at: <https://doi.org/10.1037/0033-2909.86.2.420>.

Sousa Santos, B. de (2007) “Beyond abyssal thinking: From global lines to ecologies of knowledges,” *Revista Crı́tica de Ciências Sociais*, 78, pp. 3–46. Available at: <https://ces.uc.pt/publicacoes/santos_beyond_abyssal_thinking.pdf>.

Spearman, C. (1904) “The proof and measurement of association between two things,” *The American Journal of Psychology*, 15(1), pp. 72–101. Available at: <https://doi.org/10.2307/1412159>.

Spyros, A., Dais, P., *et al.* (2023) “Current state-of-the-art spectroscopic and chromatographic techniques utilized in food authenticity and food traceability,” *Foods*, 13(1), p. 3. Available at: <https://doi.org/10.3390/foods13010003>.

Streiner, D.L. and Norman, G.R. (2008) *Health measurement scales: A practical guide to their development and use*. 4th ed. Oxford: Oxford University Press.

Student (1908) “The probable error of a mean,” *Biometrika*, 6(1), pp. 1–25. Available at: <https://doi.org/10.2307/2331554>.

Surname, A. and Surname, A. (2025) “A comparative study of 3-point and 5-point likert scales,” *Science Scholar* [Preprint].

Team, R.C. (2024) *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <https://www.R-project.org/>.

Thompson, S.G. and Higgins, J.P.T. (2002) “How should meta-regression analyses be undertaken and interpreted?” *Statistics in Medicine*, 21(11), pp. 1559–1573. Available at: <https://doi.org/10.1002/sim.1187>.

Torchiano, M. (2020) *Effsize: Efficient effect size computation*. Available at: <https://doi.org/10.5281/zenodo.1480624>.

Tranfield, D., Denyer, D. and Smart, P. (2003) “Towards a methodology for developing evidence-informed management knowledge by means of systematic review,” *British Journal of Management*, 14(3), pp. 207–222.

Tricco, A.C. *et al.* (2018) “PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation,” *Annals of Internal Medicine*, 169(7), pp. 467–473.

Tricco, A.C., Lillie, E., *et al.* (2018) “PRISMA extension for scoping reviews (PRISMAScR): Checklist and explanation,” *Annals of Internal Medicine*, 169(7), pp. 467–473. Available at: <https://doi.org/10.7326/M18-0850>.

Urvieta, R. *et al.* (2021) “Terroir and vintage discrimination of malbec wines based on phenolic composition across multiple sites in mendoza, argentina,” *Scientific Reports*, 11(1). Available at: <https://doi.org/10.1038/s41598-021-82306-0>.

Vandecandelaere, E. *et al.* (2009) *Linking people, places and products: A guide for promoting quality linked to geographical origin and sustainable geographical indications*. Rome: FAO; SINER-GI.

Viechtbauer, W. (2010) “Conducting meta-analyses in r with the metafor package,” *Journal of Statistical Software*, 36(3), pp. 1–48. Available at: <https://doi.org/10.18637/jss.v036.i03>.

Wadoux, A.M.J.C. *et al.* (2021) “Spatial cross-validation is not the right way to evaluate map accuracy,” *Ecological Modelling*, 457, p. 109692. Available at: <https://doi.org/10.1016/j.ecolmodel.2021.109692>.

Wang, X. *et al.* (2025) “Integrated metabolomics-KPCA-machine learning framework: A solution for geographical traceability of chinese jujube,” *Food Chemistry: X*, 31. Available at: <https://doi.org/10.1016/j.fochx.2025.103069>.

Webster, J. and Watson, R.T. (2002) “Analyzing the past to prepare for the future: Writing a literature review,” *MIS Quarterly*, 26(2), pp. xiii–xxiii.

Wickham, H. (2016) *ggplot2: Elegant graphics for data analysis*. Springer. Available at: <https://ggplot2.tidyverse.org>.

Wilkinson, L. (2021) “The grammar of graphics,” *Springer Series in Statistics* [Preprint]. 2nd ed.

Wilkinson, M.D. *et al.* (2016) “The FAIR guiding principles for scientific data management and stewardship,” *Scientific Data*, 3, p. 160018. Available at: <https://doi.org/10.1038/sdata.2016.18>.

World Bank (2014) “Open data for agriculture and nutrition,” *World Bank Report* [Preprint]. Available at: <https://www.worldbank.org/en/topic/agriculture/brief/open-data-for-agriculture>.

World Intellectual Property Organization (2020) *World intellectual property indicators 2020*. Geneva: WIPO. Available at: <https://doi.org/10.34667/tind.42184>.

World Trade Organization (1994) “Agreement on trade-related aspects of intellectual property rights (TRIPS agreement).” Marrakesh. Available at: <https://www.wto.org/english/docs_e/legal_e/27-trips.pdf>.

Xu, F. *et al.* (2021) “Combing machine learning and elemental profiling for geographical authentication of chinese geographical indication (GI) rice,” *npj Science of Food*, 5(1), p. 18. Available at: <https://doi.org/10.1038/s41538-021-00100-8>.

Young, I.J.B., Luz, S. and Lone, N. (2019) “A systematic review of natural language processing for classification tasks in the field of incident reporting and adverse event analysis,” *International Journal of Medical Informatics*, 132, p. 103971. Available at: <https://doi.org/10.1016/j.ijmedinf.2019.08.018>.