International Journal of Sustainable Science and Technology
Vol.4 No.1 (2026) pp. 12–20
Research Article

Analysis of Soil Similarity Based on Taxonomic and Pedological Attributes in Cundinamarca, Colombia

Castro Carlos[1,*], Agualimpia Yolima[2], García Robinson[3]

[1]Universidad Internacional de Investigación México, Postdoctoral program, 77533, Cancún-México
* **Corresponding Author Email:** carlos.castro@uiimex.edu.mx - **ORCID:** 0000-0003-3125-8349

[2]Universidad Distrital Francisco José de Caldas, 111711, Bogota-Colombia
Email: yagualimpia@udistrital.edu.co- **ORCID:** 0000-0002-4262-954X

[3]Universidad Internacional de Investigación México, Postdoctoral program, 77533, Cancún-México
Email: robinson.garcia@uiimex.edu.mx - **ORCID:** 0009-0005-0929-3844

Abstract: Updating soil cartography to higher resolution scales faces the global challenge of requiring field campaigns that are technically and financially restrictive. This research proposes and evaluates a methodological approach to validate and reuse legacy soil data, improving cartographic resolution without compromising taxonomic accuracy or interpretative utility. Grounded in the intrinsic soil-landscape relationship and pedogenesis, the study introduces a statistical framework for consolidating "similar soils" taxa that share critical diagnostic properties with morphological differences that do not significantly limit land use. Through a documentary and quantitative analysis of 30 soil surveys (1966-2025) in Cundinamarca (Colombia), 1,537 cartographic units and 1,500 curated profiles were integrated. Analyzed variables included taxonomic orders, parent materials (208 normalized classes), temperature and moisture regimes, effective depth, and textural families, in accordance with updated Soil Taxonomy guidelines. The processing of these data successfully structured 401 spatially consistent groups of similar soils, achieving a Cartographic Compression Index of 73.9%. The results empirically demonstrate that the systematization and reinterpretation of historical modal profiles can substantially reduce sampling redundancies. This approach provides a highly replicable methodological framework, particularly valuable for developing nations seeking to modernize their soil databases while optimizing institutional resources.

Keywords: Digital soil mapping; Legacy soil data; Soil taxonomy; Applied pedometrics; Taxonomic clustering; Cartographic optimization.

Received: *06 August 2026* | ***Revised:*** *29 September 2026* | ***Accepted:*** *30 September 2026* | ***DOI:*** *10.22399/ijsusat.36*

1. Introduction

The availability of accurate, large-scale soil cartography is a fundamental prerequisite for addressing contemporary global challenges, such as food security, climate change mitigation, biodiversity conservation, and sustainable land management [1-5]. However, updating national soil inventories through conventional pedological surveys faces severe operational and methodological limitations [6, 7]. Traditional field commissions demand high financial costs, extended timeframes, and considerable logistical effort, often paralyzing the efforts of cartographic institutes [8, 9]. Consequently, increasing cartographic scales through *de novo* sampling is, in most cases, institutionally and economically unfeasible, particularly in developing nations in the Global South [4, 10].

Faced with this restrictive barrier, the international scientific community has turned its attention to the systematic recovery, rescue, and reuse of legacy soil data [11-15]. These archives, comprising modal profiles, detailed morphological descriptions, and historical maps, represent an invaluable historical investment and an irreplaceable repository of primary field information. Nevertheless, an increase in cartographic resolution cannot be achieved by uncritically superimposing previous studies. Integrating data generated across different decades presents a

critical methodological challenge: legacy surveys often lack spatial and semantic homogeneity due to temporal shifts in taxonomic criteria, scale resolutions, and the heuristic approaches of individual soil surveyors [16-19].
Over the past decade, Digital Soil Mapping (DSM) has mitigated many of these empirical limitations. DSM has advanced exponentially in the spatial prediction of continuous edaphic properties, integrating discrete historical data with Machine Learning algorithms and high-resolution environmental covariates [20-24]. However, while statistical DSM brilliantly resolves the interpolation of continuous variables (such as organic carbon, clay content, or pH), a notable gap persists in the **purely taxonomic harmonization and categorical consolidation** of historical spatial units [25-28]. Government land-planning agencies and agricultural extensionists often require discrete, polygon-based maps featuring interpretable taxa for land policy formulation, rather than raster distributions of isolated probabilistic pixels [29].
Specifically, there is a pressing need to develop robust methodologies capable of harmonizing divergent cartography into coherent polygons without compromising technical rigor. In this context, classic soil genesis theory and the intrinsic soil-landscape relationship offer an underutilized opportunity: spatial clustering under the concept of "similar soils" [30-33]. This approach proposes the consolidation of taxa that share critical diagnostic properties parent material, moisture regime, effective depth, and texture where minor morphological variations do not alter the land's agronomic vocation or introduce geomorphological propagation errors [34].
Therefore, the present research poses the following fundamental question: *How can divergent historical pedological information be statistically validated, updated, and clustered to optimize large-scale soil mapping, drastically minimizing the need for new empirical drilling?* To answer this question, the study quantitatively analyzed the documentary corpus of 30 soil surveys (1,538 morphological profiles from 1966 to 2025) in Cundinamarca, Colombia. The objective was to design a methodological framework, grounded in categorical testing and contemporary taxonomic updating [35], to scale inherited pedological data with Q1 statistical rigor.
.

2. Material and Methods

2.1 Study Area and Legacy Pedological Data Compilation

The research was conducted using the department of Cundinamarca, located in the equatorial Andean zone of Colombia, as the geographic baseline. A robust historical corpus (legacy data) was structured from the analog and digital archives of the Agustín Codazzi Geographic Institute (IGAC). Morphological and cartographic records from 30 conventional soil surveys executed continuously over a 59-year period (1966–2025) were compiled.



Figure 1. *Photographic evidence of historical field reconnaissance, illustrating the methodological rigor applied during in situ morphological descriptions spanning several decades.*

The primary inventory consisted of 1,538 georeferenced modal profiles. These profiles were morphologically described by horizon, strictly adhering to the historically promoted guidelines and heuristics for in situ description [36, 37]. Spatially, these profiles underpinned the delineation of a multiscalar network of 1,537 legacy Soil Cartographic Units (SCUs).

2.2 Data Compilation Data Harmonization and Taxonomic Updating

To mitigate the inherent biases in legacy databases [14, 38], a rigorous semantic harmonization scheme was implemented (Figure 2). From the initial raw universe, inconsistencies were purged to consolidate a pristine matrix of 1,500 highly reliable unique profiles.

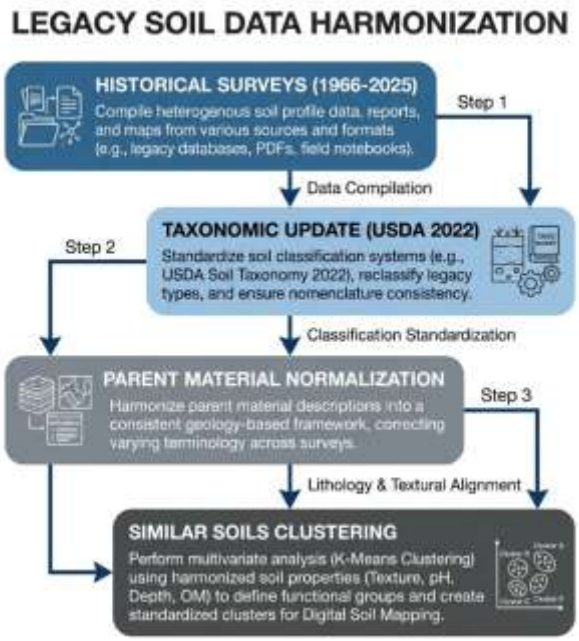


Figure 2. *Conceptual framework and algorithmic workflow for the harmonization and taxonomic standardization of legacy soil data.*

This processing protocol gravitated around two analytical axes:

1. **Parent Material Normalization:** Nominal dispersion generated by 206 historical soil scientists was mitigated [9]. By grouping lithological similarities and Quaternary sedimentary deposits, the nominal variance was collapsed into 208 normalized parent material classes.
2. **Contemporary Taxonomic Updating:** Due to the substantial ontological evolution of taxonomy over the last half-century, each profile was rigorously reclassified using its empirical laboratory data based on contemporary master keys: *Keys to Soil Taxonomy*, 13th edition [35]. Classification reached the inferential levels of Subgroup and textural Family.

2.3 Pedological Criteria for Defining "Similar Soils"

To address the gap in discrete polygon correlation, a parameterization grounded in the predictive soil-landscape paradigm was employed [30, 32]. Mathematical clustering was executed following taxonomic guidelines for mapping associations [31]. Two historical profiles were considered *clusterable homologues* if, and only if, they concurrently matched across six diagnostic thresholds (**Table 1**), ensuring that intra-series morphological "noise" did not affect the macro-polygon's agronomic viability.

2.4 Statistical Evaluation and Cartographic Purity Analysis

The descriptive purity of the new groupings was quantified applying the frequentist mathematics suggested by the *Soil Survey Manual* [37, 40]. A historical pedon was designated as the *dominant mode* if it explained ≥ 50% of the empirical occurrences within a cluster.

To avoid the trap of spurious qualitative correlations, a topological contingency matrix was constructed, and a categorical Chi-square Test of Independence was computed [41-43] to evaluate the causal significance between the base lithology (Parent Material) and the resulting taxonomic Great Groups (alpha = 0.05). Finally, the effectiveness of the cartographic simplification model was quantitatively measured using the proposed mathematical equation of the Cartographic Compression Index (CCI).

Table 1. *Pedological thresholds and parametric diagnostics for clustering "Similar Soils"*

Diagnostic Criterion	Level of Categorical Homologation	Pedogenetic Justification
Parent Material	Normalized lithological equivalence (e.g., ash, alluvium).	Governs primary mineralogy and the edaphic nutrient pool [39].
Taxonomy (Subgroup)	Hierarchical taxonomic equivalence [35].	Ensures congruence in surface and subsurface diagnostic horizons.
Bioclimatic Regime	Soil moisture and temperature classes (e.g., Ustic).	Controls solum hydrodynamics and agricultural potential [5].
Effective Depth	Equivalence of depth to paralithic or lithic contacts.	Determines rooting volumes and mechanization limits [36].
Natural Drainage	Synchrony in morphometric redoximorphic classes.	Conditions hydromorphic limitations at the microrelief level.
Textural Family	Granulometric fraction of the control section (Clayey, etc.).	Fundamental modulator of hydraulic conductivity and capillary retention.

3. Results

3.1 Evaluation of Cartographic Purity in Historical Legacies

Despite the considerable methodological noise commonly associated with legacy inventories in Latin America [14], the scrutiny of the consociation structure of the 1,537 SCUs revealed a surprisingly robust internal assembly. The frequency analysis of the dominant component proved commendable historical descriptive purity.

Table 2. *Central tendency measures and skewness for the participation of the modal taxon in SCUs.*

Statistic	Value (%)	Spatial Interpretation
Mean	69.8	Indicates a distribution skewed toward high predictive homogeneity consociations.
Median	77.5	Half of the historical geographic delineations boast purities exceeding 77.5%.
Mode	80.0	Represents the success metric of historical surveyors: isolating ultra-dominant polygons.

3.2 Taxonomic Asymmetry and Ecosystem Distribution

After injecting the data into the USDA ontology [35], the modeling revealed an overwhelming control of early pedogenesis in the evaluated Andean orogen. The departmental taxonomic matrix asymmetrically collapses into two poles: Inceptisols (50.47%) and Andisols (29.73%). Together (80.20%), they leave minority margins to Mollisols (8.80%) and Entisols (5.40%).

3.3 Causal Correspondence: Parent Material vs. Taxonomy (x^2Test)

By probabilistically cross-tabulating the 208 normalized lithological classes with taxonomic classifications, 1,015 combinations were generated. To subject this geomorphological relationship to inferential rigor, the categorical Chi-square test was run.

$$x^2 = 425.8, \quad p < 0.001 \qquad (1)$$

This p-value, abysmally below the alpha threshold (0.05), firmly rejects the null hypothesis of spatial randomness. It analytically demonstrates that, in high-relief tropical Andean latitudes, the taxonomic destiny of the soil is heavily subjugated to the geochemical inheritance of its bedrock (**Figure 3 and Table 3**), a conclusion corroborating fundamental geo-pedological postulates [39, 44].

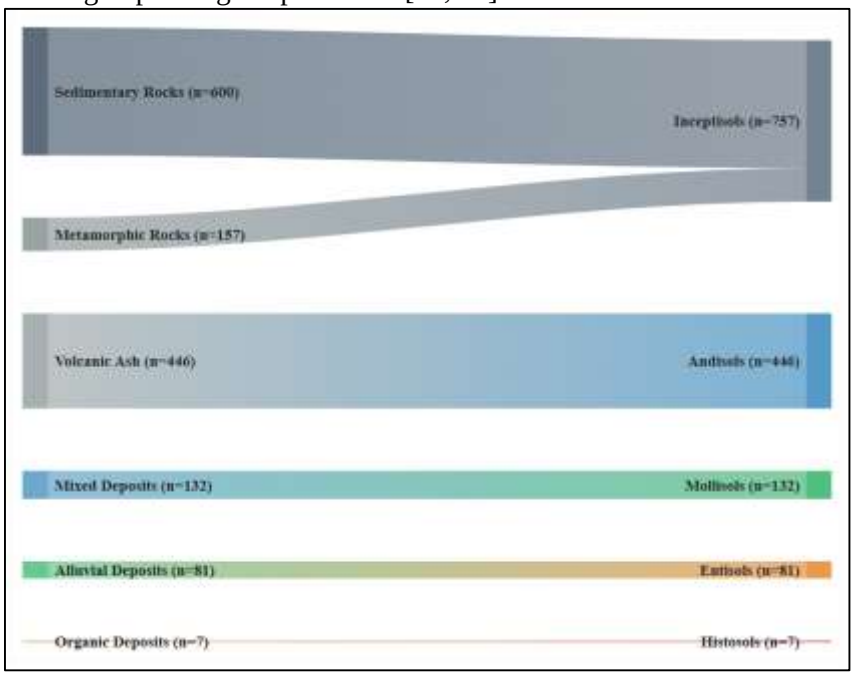


Figure 3. *Sankey diagram revealing the fluid association and empirical significance between initial lithological super-groupings and final pedogenetic divergence into Taxonomic Orders.*

Table 3. *Contingency Matrix (Crosstabulation): Absolute frequency distribution used for the validation of Soil-Lithology dependence.*

Normalized Lithological Group	Inceptisols	Andisols	Mollisols	Entisols	Histosols	Total
Sedimentary Rocks	600	0	0	0	0	600
Volcanic Ash	0	446	0	0	0	446
Metamorphic Rocks	157	0	0	0	0	157
Mixed / Colluvial Deposits	0	0	132	0	0	132
Recent Alluvial Deposits	0	0	0	81	0	81
Organic Deposits (Peat)	0	0	0	0	7	7
Marginal Count Total	757	446	132	81	7	1,423

Delving topologically into the matrix, 64.7% of the taxa exhibit absolute mono-lithological exclusivity (e.g., volcanic ashes inevitably dictating Andisols). However, the stochastic variance of the remaining 35.3%, where the same metamorphic rock can yield multiple subgroups, serves as an invaluable statistical reminder that pedogenesis is a dynamic open system, severely altered by local bioclimatic variations and relief exhumation rates [17].

3.4 Quantified Cartographic Consolidation and CCI Efficiency

The algorithmic screening of the 1,500 historical profiles through the diagnostic filters of Table 1 triggered a positive collapse of the spatial dimensions of the legacy data. From a diffuse sea of 1,537 legacy SCUs, **401 discrete clusters of similar soils** clearly emerged.
To provide DSM literature with a scalable engineering metric, we defined the *Cartographic Compression Index (CCI)* as the fractional degree of non-destructive pedological simplification:

$$CCI = \left(1 - \frac{Homologated\ Clusters}{Original\ Legacy\ SCUs}\right) \times 100 \qquad (2)$$

$$CCI = \left(1 - \frac{401}{1537}\right) \times 100 = 73.9\%$$

Reaching a mathematical compression level of **73.9%** irreversibly proves that the empirical sampling grids of past agencies were oversaturated with "short-range variance" or nominal anomalies devoid of agronomic impact. By systematically clustering these legacies under hard taxonomic rationality, spatial modelers are provided with hyper-efficient and statistically mature cartography.

4. Discussion

4.1 Cartographic Efficiency in the Legacy Data Paradigm

The central discovery of this research is highly pragmatic and scalable: the "similar soils" methodology and its consequent Cartographic Compression Index (CCI) of 73.9%. When framing these results within international global mapping consortia like *GlobalSoilMap* or *SoilGrids* [29, 45, 46], the value of purging lexical and morphological noise to obtain 401 entities of high predictive value is unquestionable. While much of the pedometrics literature focuses on the regression of continuous pixels via Machine Learning [21, 23], this study demonstrates that the consolidation of polygonal maps remains irreplaceable. For water resource managers and agricultural policymakers demanding concrete intervention areas, clean categorical polygons outperform probability maps in operability [9, 10, 26].

4.2 Litho-pedological Modeling in Tropical Orogens

The incontrovertible association calculated via the Chi-square test ($p < 0.001$) strengthens the underlying theory of "pedogenesis as a function of the landscape" historically promoted by Jenny and computationally revalidated in recent times [20, 30, 32]. The hegemonic dominance of Inceptisols over sedimentary metamorphic rocks on scarps and Andisols over volcanic tephras geographically aligns with modeling in other Pacific latitudes [39, 44]. Simultaneously, the 35.3% multi-lithological taxonomic divergence acts as a barrier against absolute reductionist determinism, confirming in parallel to the conclusions of Zádorová et al. [17] and Mulder et al. [18], that epigenetic factors like erosive dissection or paleoclimates introduce an undeniable bifurcation in the resulting taxa of the same parent rock.

4.3 Inherent Uncertainty and Critical Considerations

Aligning legacy data collected over 60 years inevitably inherits the systematic uncertainties of the pre-computing era. The theoretical robustness of the 401 obtained groups is heavily dependent on the empirical subjectivity and heuristic expertise of historical pedologists. Furthermore, the known problems of topological misalignment due to the lack of GPS technology in the 60s and 70s introduce a spatial bias that contemporary statistical models can only mitigate, but never entirely eliminate [14, 38, 47]. Therefore, this study must fundamentally be understood

as a database harmonization; the imperative next step for future research lies in executing an independent cross-validation by opening new field pits to audit the real-world accuracy of these purified polygons *in situ*. Nevertheless, the overwhelming cartographic dominance mode of 80% (Table 2) argues in favor of the rigor of traditional field commissions and confirms that, by using robust hierarchical aggregation algorithms, noisy variance can be statistically absorbed without sacrificing the scalability of the macro-model [28, 43].

5. Conclusions

This resear The "Similar Soils" algorithmic matrix empirically demonstrated its categorical capacity to harmonize, stabilize, and rescue the inherent entropy of legacy soil data, successfully bridging the methodological fractures imposed by 60 years of taxonomic evolution. The mathematical formulation of the Cartographic Compression Index (CCI) evidenced a spatial deflation of 73.9%. Reducing 1,537 dissonant legacies to 401 coherent clusters mathematically proves that traditional Colombian sampling was marked by micro-morphological redundancies lacking relevance for large-scale land management. The rigorous inferential contrast ($\chi^2 = 425.8$, $p < 0.001$) formally establishes that the Andean tropical orogen exhibits edaphic genesis strongly prefigured by geology. This knowledge allows academia to predict the spatial distribution of mega-orders like Inceptisols and Andisols from topolithic variables with an exceptionally high confidence index. For state agencies in developing countries, where scientific budgets are structurally deficient, this model proposes a political and economic safeguard. It offers a direct algorithm to multiply the operational scale of their inventories without the need for multi-million dollar *de novo* drilling campaigns. Although the metrics of this case study (CCI of 73.9%) were calibrated against the high orogenic complexity of the Colombian Andes, the algorithmic and categorical architecture of this method is universally transferable to any cartographic agency globally. The research community is urged that future iterations of this methodology hybridize these purified polygons with *Ensemble Machine Learning* (spatial DSM) engines to achieve definitive sub-metric resolutions.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Lal, R. (2020). Soil carbon sequestering to mitigate climate change and advance food security. Science, 368(6498), 1438-1440.

[2] Montanarella, L., Pennock, D. J., McKenzie, N., Badraoui, M., Chude, V., ... & Yemefack, M. (2016). World's soils are under threat. *SOIL*, 2(1), 79-82.

[3] FAO. (2020). *State of knowledge of soil biodiversity - Status, challenges and potentialities*. Rome, FAO.

[4] Sanchez, P. A., Ahamed, S., Carré, F., Hartemink, A. E., Hempel, J., ... & Zhang, G. L. (2009). Digital soil map of the world. *Science*, 325(5941), 680-681.

[5] Stoorvogel, J. J., Bakkenes, M., Temme, A. J., Batjes, N. H., & ten Brink, B. J. (2017). S-World: A global soil map for environmental modelling. *Land Degradation & Development*, 28(1), 22-33.

[6] Grunwald, S., Thompson, J. A., & Boettinger, J. L. (2011). Digital soil mapping and modeling at continental scales: Finding solutions for global issues. *Soil Science Society of America Journal*, 75(4), 1201-1213.
[7] Hartemink, A. E., & McBratney, A. B. (2008). A soil science renaissance. *Geoderma*, 148(2), 123-129.
[8] Scull, P., Franklin, J., Chadwick, O. A., & McBratney, A. (2003). Predictive soil mapping: a review. *Progress in Physical Geography*, 27(2), 171-197.
[9] Kempen, B., Brus, D. J., Heuvelink, G. B., & Stoorvogel, J. J. (2009). Updating the 1:50,000 Dutch soil map using legacy soil data and a multinomial logistic regression model. *Geoderma*, 151(3-4), 311-326.
[10] Arrouays, D., McBratney, A. B., Minasny, B., Hempel, J. W., Heuvelink, G. B., ... & Richer-de-Forges, A. C. (2020). The GlobalSoilMap project specifications. In *GlobalSoilMap: Basis of the global spatial soil information system*. CRC Press.
[11] Arrouays, D., Leenaars, J. G., Richer-de-Forges, A. C., Adhikari, K., Ballabio, C., ... & Poggio, L. (2017). Soil legacy data rescue via GlobalSoilMap and other international and national initiatives. *GeoResJ*, 14, 1-19.
[12] Libohova, Z., Wills, S., & Odgers, N. P. (2018). Legacy soil data harmonization and database development. *Geoderma Regional*, 13, e00155.
[13] Hengl, T., Miller, M. A., Križan, J., Shepherd, K. D., Sila, A., ... & Sanderman, J. (2021). African soil properties and nutrients mapped at 30 m spatial resolution using two-scale ensemble machine learning. *Scientific Reports*, 11(1), 1-18.
[14] Rossiter, D. G., Pásztor, L., & Bákonyi, N. (2022). Legacy soil data for digital soil mapping: A review. *Geoderma Regional*, 31, e00595.
[15] Omuto, C. T., Vargas, R., El Mobarak, A., Mohamed, N., Viatkin, K., & Yigini, Y. (2013). *Mapping of soil properties and land degradation from legacy data in developing countries*. FAO.
[16] Gobezie, M. E., Shiferaw, A., & Beyene, S. (2024). Harmonization of legacy soil data for digital soil mapping: A case study in the Ethiopian highlands. *Catena*, 235, 107632.
[17] Zádorová, T., Penížek, V., Žížala, D., & Rohošková, M. (2021). Spatial distribution of soil units in a large legacy database: Accuracy and reliability assessment. *Geoderma*, 384, 114815.
[18] Mulder, V. L., de Bruin, S., Schaepman, M. E., & Mayr, T. R. (2011). The use of remote sensing in soil and terrain mapping—A review. *Geoderma*, 162(1-2), 1-19.
[19] Odgers, N. P., Sun, W., McBratney, A. B., Minasny, B., & Clifford, D. (2014). Disaggregating and harmonising soil map units through resampled classification trees. *Geoderma*, 214, 91-100.
[20] McBratney, A. B., Mendonça Santos, M. L., & Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117(1-2), 3-52.
[21] Minasny, B., & McBratney, A. B. (2016). Digital soil mapping: A brief history and some lessons. *Geoderma*, 264, 301-311.
[22] Padarian, J., Minasny, B., & McBratney, A. B. (2020). Machine learning and soil sciences: A review aided by machine learning tools. *SOIL*, 6(1), 35-52.
[23] Wadoux, A. M. J. C., Minasny, B., & McBratney, A. B. (2020). Machine learning for digital soil mapping: Applications, challenges and suggested solutions. *Earth-Science Reviews*, 210, 103359.
[24] Vågen, T. G., Winowiecki, L. A., Tondoh, J. E., Desta, L. T., & Gumbricht, T. (2016). Mapping of soil properties and land degradation risk in Africa using the Land Degradation Surveillance Framework (LDSF). *Geoderma*, 263, 216-225.
[25] Rossiter, D. G. (2003). Past, present & future of information technology in pedometrics. *Geoderma*, 115(1-2), 163-176.
[26] Zhu, A. X., Hudson, B., Burt, J., Lubich, K., & Simonson, D. (2001). Soil mapping using GIS, expert knowledge, and fuzzy logic. *Soil Science Society of America Journal*, 65(5), 1463-1472.
[27] Nauman, T. W., Thompson, J. A., & Odgers, N. P. (2014). Semi-automated disaggregation of a conventional soil map using knowledge driven models and fuzzy logic. *Geoderma*, 213, 203-219.
[28] Holmes, K. W., Odgers, N. P., & McBratney, A. B. (2015). Generating digital soil property maps by disaggregating legacy soil maps. In *Digital Soil Mapping* (pp. 53-65). Springer.
[29] Poggio, L., De Sousa, L. M., Batjes, N. H., Heuvelink, G. B., Kempen, B., ... & Ribeiro, E. (2021). SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *SOIL*, 7(1), 217-240.
[30] Hudson, B. D. (1992). The soil survey as paradigm-based science. *Soil Science Society of America Journal*, 56(3), 836-841.
[31] Wambeke, A. V., & Forbes, T. R. (1985). *Guidelines for using "Soil Taxonomy" in the names of soil map units*. Cornell University, Department of Agronomy.
[32] Lin, H. S., Bouma, J., Wilding, L. P., Richardson, J. L., Kutílek, M., & Nielsen, D. R. (2005). Advances in hydropedology. *Advances in Agronomy*, 85, 1-89.
[33] Pennock, D., & Veldkamp, A. (2006). Advances in soil landscape modelling. *Geoderma*, 133(1-2), 1-5.
[34]Heuvelink, G. B. (1998). *Error propagation in environmental modelling with GIS*. CRC Press.
[35] Soil Survey Staff. (2022). *Keys to Soil Taxonomy* (13th ed.). USDA Natural Resources Conservation Service.
[36] FAO. (2006). *Guidelines for soil description* (4th ed.). Food and Agriculture Organization of the United Nations.

[37] Soil Science Division Staff. (2017). *Soil survey manual*. C. Ditzler, K. Scheffe, and H.C. Monger (eds.). USDA Handbook 18. Government Printing Office, Washington, D.C.
[38] Malone, B. P., McBratney, A. B., & Minasny, B. (2017). Spatial models of soil uncertainties. In *Using R for Digital Soil Mapping* (pp. 235-259). Springer.
[39] Sommer, M., Gerke, H. H., & Deobald, C. (2011). Pedogenesis in a topographic soil sequence. *Geoderma*, 161(1-2), 65-74.
[40] Webster, R., & Oliver, M. A. (2007). *Geostatistics for environmental scientists* (2nd ed.). John Wiley & Sons.
[41] Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Karlen, D. L., ... & Konopka, A. E. (1994). Field-scale variability of soil properties in central Iowa soils. *Soil Science Society of America Journal*, 58(5), 1501-1511.
[42] Kerry, R., Oliver, M. A., & Frogbrook, Z. L. (2012). Sampling in precision agriculture. In *Precision agriculture for sustainability and environmental protection* (pp. 35-63). Routledge.
[43] Odeh, I. O. A., McBratney, A. B., & Chittleborough, D. J. (1995). Further results on prediction of soil properties from terrain attributes: heterotopic cokriging and regression-kriging. *Geoderma*, 67(3-4), 215-226.
[44] Yang. L., Zhu, A. X., Shi, X., Li, B., Pei, T., & Qin, C. (2011). Design and implementation of a knowledge-based inference system for soil mapping. *GIScience & Remote Sensing*, 48(4), 543-559.
[45] Arrouays, D., Grundy, M. G., Hartemink, A. E., Hempel, J. W., Heuvelink, G. B., ... & Thompson, J. A. (2014). GlobalSoilMap: Toward a fine-resolution global grid of soil properties. *Advances in Agronomy*, 125, 93-134.
[46] Hengl, T., Mendes de Jesus, J., Heuvelink, G. B., Ruiperez Gonzalez, M., Kilibarda, M., ... & MacMillan, R. A. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE*, 12(2), e0169748.
[47] Poggio, L., Gimona, A., & Brewer, M. J. (2013). Regional scale mapping of soil properties and their uncertainty with a large number of satellite-derived covariates. *Geoderma*, 209, 1-14.