Eurasian Journal of Soil Science 2026, 15(1), 141 - 148

Eurasian Journal of Soil Science

Journal homepage : http://ejss.fesss.org

Theoretical foundations for developing a digital soil twin for Southern Russia

Yurii Litvinov [a,*], Anton Mezhenkov [a], Tanwar Sudip [a], Svetlana Sushkova [a], Alexey Samoylov [b], Eduard Melnik [b], Alexander Kozlovskiy [b], Ekaterina Kuchmenko [a], Tatiana Minkina [a], Evgenyi Shuvaev [a], Anastasia Nemtseva [a], Rahila İslamzade [c,d]

[a] Academy of Biology and Biotechnology, Southern Federal University, Rostov-on-Don, Russia
[b] Institute of Computer Technologies and Information Security, Southern Federal University, Taganrog, Russia
[c] Azerbaijan State Oil and Industry University, Baku, Azerbaijan
[d] Sumgayit State University, Sumgayit, Azerbaijan

Article Info
Received : 03.05.2025
Accepted : 08.12.2025
Available online: 12.12.2025

Author(s)
Y.Litvinov *
AMezhenkov
T.Sudip
S.Sushkova
A.Samoylov
E.Melnik
A.Kozlovskiy
E.Kuchmenko
T.Minkina
E.Shuvaev
A.Nemtseva
R.İslamzade

* Corresponding author

Abstract

The sustainable management of soil resources in Southern Russia has become increasingly critical due to intensifying agricultural pressures, accelerating climate variability, and cumulative anthropogenic disturbances. Rapid advances in digital technologies now enable the integration of heterogeneous soil datasets into dynamic computational environments capable of representing physical soil systems with unprecedented precision. Within this context, the digital twin (DT) paradigm—originating from engineering sciences and now rapidly expanding into agricultural and environmental domains—offers a transformational framework for real-time soil monitoring, process simulation, scenario forecasting, and risk assessment. This study establishes the theoretical and methodological foundations necessary for developing a comprehensive soil digital twin for the Rostov region by synthesizing more than eighty years of archival soil–geographical surveys, long-term agrochemical monitoring data, multi-scale cartographic sources, remote sensing products, IoT-based soil measurements, climate records, machine-learning algorithms, geostatistical models, semantic graph structures, distributed computing frameworks, federated learning, and blockchain-enabled data governance. Particular emphasis is placed on harmonizing heterogeneous soil legends, vectorizing analog soil maps, constructing unified soil ontologies, and designing a multi-layered DT architecture grounded in contemporary digital twin theory, including mirrored physical–virtual spaces, multidimensional modeling, and state-fusion mechanisms. Machine-learning experiments demonstrate high predictive accuracy for numerous soil attributes, while geostatistical modeling enhances spatial continuity and uncertainty quantification. The integrated framework presented here provides a robust foundation for constructing an operational soil digital twin capable of supporting precision agriculture, environmental monitoring, insurance modeling, and strategic land-use planning. By enabling continuous data ingestion, multi-stakeholder interaction, and dynamic model refinement, the developed digital twin concept has significant potential to strengthen climate resilience, optimize agronomic interventions, and promote sustainable agricultural development across Southern Russia.

Keywords: Digital twin, soil, Rostov region, modeling, GIS, IoT, big data, artificial intelligence.

© 2026 Federation of Eurasian Soil Science Societies.

Introduction

The sustainable management, monitoring, and long-term preservation of soil resources have become increasingly urgent globally, and the situation is especially acute in Southern Russia, where some of the

: https://doi.org/10.18393/ejss.1841018
: https://ejss.fesss.org/10.18393/ejss.1841018
Publisher : Federation of Eurasian Soil Science Societies
e-ISSN : 2147-4249

world's most productive chernozem soils coexist with intensive agricultural exploitation and growing climatic instability. The region faces simultaneous pressures from drought, erosion, salinization, soil compaction, nutrient depletion, and anthropogenic disturbances that threaten agronomic productivity and ecosystem resilience. Addressing these interacting risks requires analytical tools capable of integrating heterogeneous multidecadal soil datasets, representing soil processes dynamically, and generating predictive insights to support evidence-based decision-making across multiple sectors. The rapid development of digital technologies—geospatial analytics, remote sensing, cloud and edge computing, sensor networks, machine learning, and artificial intelligence—has made it possible to conceptualize soil systems not merely as static datasets but as dynamic, data-driven, computationally accessible entities.

Against this backdrop, the digital twin (DT) paradigm has emerged as a transformative framework for modeling and managing complex agricultural and ecological systems. Originally conceptualized in aerospace and manufacturing engineering, modern DT theory extends far beyond static simulation, incorporating continuously synchronized virtual representations, behavioral modeling, and multi-source state fusion. Recent foundational research describes the DT as a multi-dimensional cyber–physical–social construct characterized by mirrored physical and virtual spaces, linkage mechanisms, rule-based behavior, and iterative cycles of construction, assembly, fusion, prediction, verification, modification, and lifecycle management (Tao et al., 2022). This multi-layered theoretical framework provides a robust foundation for representing highly heterogeneous, nonlinear, and dynamic environmental systems such as soils.

Parallel advances in agricultural digitalization further strengthen the feasibility of implementing soil digital twins. Studies demonstrate that integrating IoT-based soil monitoring systems, real-time sensor networks, autonomous data collection, and cloud-based analytics can significantly enhance the precision and timeliness of agricultural decision-making. These systems produce dense streams of information on soil moisture, temperature, salinity, and nutrient dynamics, enabling real-time digital representations of field conditions and supporting precision agriculture and resource-efficient management (Nasirahmadi and Hensel, 2022).

Moreover, recent research emphasizes the need for ecosystem-scale DTs that integrate remote sensing, ground measurements, climate data, land-use information, and simulation outputs to support sustainable land management and ecological risk assessment (Tagarakis et al., 2024). A systematic review on soil digital twins highlights that, while numerous studies have demonstrated DT applications in irrigation optimization (Borovoy et al., 2024), fertilizer management (Petroselli et al., 2024), evapotranspiration prediction (Safonova et al., 2023), soil moisture estimation (Parewai and Köppen, 2025), soil quality forecasting (Silva et al., 2023), ecosystem monitoring (Tsakiridis et al., 2023), slope stability modeling (Chen et al., 2023), and climate-related risk assessment (Gordov, 2023), most implementations remain fragmented. These systems typically address highly specialized scientific questions and lack the integrative capability to serve a broader multi-stakeholder community.

This fragmentation is particularly notable in regions such as Southern Russia, where enormous volumes of soil–geographical information have been accumulated for more than eighty years through extensive surveys conducted by YUGGIPROZEM and Southern Federal University. These archives include soil maps at multiple scales, morphological descriptions, analytical measurements, and agrochemical monitoring data collected since 1964 (FSBI SCAS, 2025), along with thematic reports on erosion, fertility, and land use. Despite their scientific value, the majority of these materials remain in analog or scanned format, with inconsistent legends, heterogeneous terminology, and varied spatial precision. As noted by Derzhavina and Bulgakova (2003) and Litvinov (2018), such datasets cannot be utilized effectively without systematic digitization, harmonization, georeferencing, and semantic standardization.

Significant progress has been made in the methodology for converting archival soil surveys into machine-readable digital resources. Vectorization frameworks (Golozubov et al., 2020), semantic harmonization techniques, unified classifiers of soil types, and methods for mapping soil combinations using weighted transparency overlays (Litvinov, 2018) have enhanced the analytical usability of historical data. The concept of elementary natural–land units (Golozubov, 2013; Bezuglova et al., 2013) further supports the integration of soil, relief, vegetation, and land-use boundaries into coherent modeling entities suitable for digital twins. These developments provide critical building blocks for constructing a regional soil DT, but they must be embedded within a more comprehensive digital architecture that integrates modern computational tools, including geostatistics, machine learning, ontology-based knowledge systems, distributed computing, and blockchain-enabled data provenance.

Digital twins provide a mechanism to merge these multiscale and multisource data environments. The incorporation of machine-learning methods enables the prediction of soil parameters that are sparsely

sampled or missing entirely from historical datasets. Prior work demonstrates that ensemble learning models—particularly Gradient Boosting and Random Forest—can effectively model nonlinear soil processes and spatial dependencies, though parameters such as exchangeable sodium and cation-exchange capacity remain more challenging due to historical sampling limitations. Geostatistical interpolation methods such as kriging and cokriging (Minasny and McBratney, 2016) complement machine learning by generating continuous spatial surfaces with quantifiable uncertainty estimates.

A fully developed soil digital twin must also support the needs of diverse users. As illustrated in Figure 1, the DT serves scientists conducting soil process modeling, agronomists managing soil fertility, government agencies responsible for land degradation monitoring, private companies testing agricultural technologies, insurance institutions assessing climate risk, farmers seeking real-time recommendations for irrigation and fertilization, and citizens contributing georeferenced environmental observations. A soil DT is therefore not merely a scientific tool but a socio-technical ecosystem that depends on interoperability, transparency, and adaptive data integration.

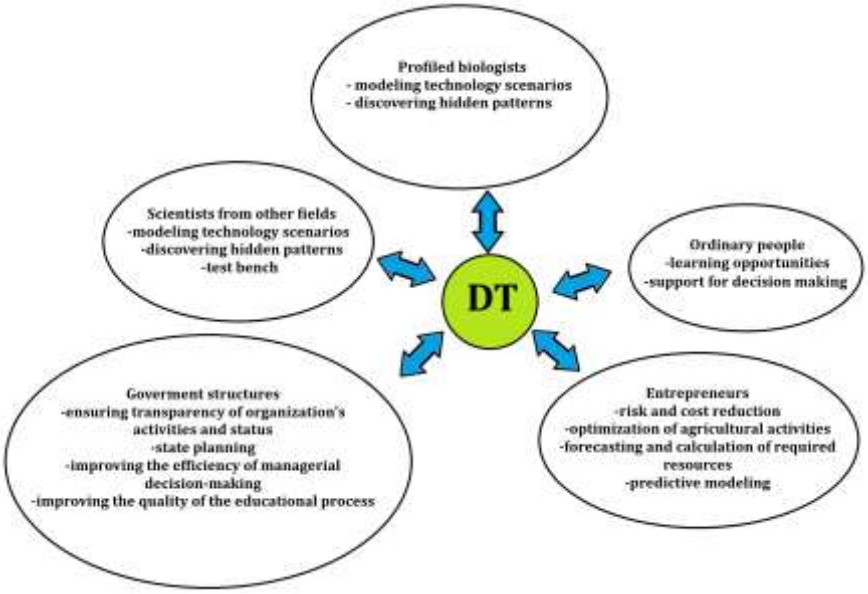


Figure 1. Stakeholder ecosystem of a digital soil twin, showing the different user groups (biologists, scientists, entrepreneurs, government bodies, and the general public) and their corresponding functional interactions with the system.

The convergence of historical soil archives, modern sensing technologies, geospatial analytics, advanced computational architectures, and contemporary DT theory creates an unprecedented opportunity to develop a comprehensive digital soil twin for the Rostov region. This study aims to establish the theoretical and methodological foundation for such a system by integrating legacy datasets, multi-source monitoring data, machine learning and geostatistical methods, semantic data structures, distributed and federated computing environments, and secure data governance frameworks. By synthesizing these components, the digital soil twin becomes a continuously evolving, analytically powerful, and operationally versatile platform capable of supporting precision agriculture, environmental monitoring, land-use planning, climate adaptation, and long-term soil sustainability in Southern Russia.

Material and Methods

The development of a regional-scale digital soil twin for the Rostov region required the integration, harmonization, and computational transformation of heterogeneous datasets spanning more than eighty years of soil–geographical observations, agrochemical monitoring programs, geospatial surveys, remote sensing products, and modern sensor-based measurements. The study area was centered on the Myasnikovsky District of the Rostov region, an agriculturally significant territory characterized by gently undulating relief, loess-derived parent materials, and highly fertile ordinary carbonate chernozems (Bezuglova, 2009). The district falls within an insufficient and unstable moisture zone (hydrothermal coefficient HTC 0.7–0.9), making it highly susceptible to drought, erosion, and salinization—conditions that necessitate the development of predictive soil models capable of simulating dynamic hydrological and agrochemical processes.

Archival soil–geographical data formed the foundational layer of the digital twin. Since 1939, YUGGIPROZEM and Southern Federal University have conducted numerous soil surveys that produced multi-scale soil maps,

profile descriptions, analytical datasets, and thematic reports on fertility, erosion, and land-use dynamics. These materials, although scientifically rich, remain largely in analog or scanned format and are characterized by multiple legend systems, non-unified classification schemes, variable coordinate precision, and inconsistent analytical methodologies. As noted in methodological manuals for soil survey and agrochemical monitoring (Derzhavina and Bulgakova, 2003; Litvinov, 2018), such heterogeneity necessitates a rigorous and multi-stage digitization process before computational modeling can be performed.

Digitization proceeded through high-resolution scanning of all cartographic materials, followed by georeferencing using an established coordinate framework. Raster soil maps were converted into vector format using the methodological principles proposed by Golozubov et al. (2020), enabling the extraction of polygon geometries, soil taxonomic units, and associated attribute information. Soil complexes and mixed units were decomposed into their constituent soil types using weighted transparency coefficients, following the approach developed by Litvinov (2018). To harmonize datasets across different survey periods, unified soil dictionaries and standardized lists of analytical indicators were constructed based on the principles of the Information System of the Soil-Geographical Database of Russia (ISSGDB). This semantic harmonization made it possible to integrate categorical, numerical, morphological, and descriptive soil information into a machine-readable, interoperable format.

To represent soil–landscape relationships within the digital twin, the methodological framework of elementary natural–land units (Golozubov, 2013; Bezuglova et al., 2013) was applied. This approach delineates homogeneous modeling entities by integrating relief forms, soil boundaries, vegetation characteristics, and land-use patterns, providing a multiscale geospatial structure suitable for predictive modeling and virtual simulations. These units formed the spatial backbone of the digital twin, enabling the connection of historical datasets with contemporary observations.

Modern data sources were incorporated to ensure continuous updating and real-time system responsiveness. IoT-based measurements from platforms such as CropX provided high-frequency observations of soil moisture, temperature, salinity, and related parameters. Satellite observations—including multispectral Sentinel-2 imagery (ESA, 2023)—contributed indices of vegetation vigor, soil brightness, and surface moisture. Climatic data were obtained from WorldClim and regional meteorological networks, while land-use information was supplemented with LUCAS statistical datasets (Eurostat, 2023). Additionally, citizen-science contributions from SoilWatch (2022) expanded spatial data coverage, particularly in under-sampled locations. All incoming data streams were processed through a dynamic ingestion pipeline that allowed both batch and real-time updates to the virtual soil environment.

Machine-learning and geostatistical methods played a critical role in estimating soil parameters that were incompletely represented in historical datasets. Using the harmonized database, predictive models were trained for a variety of soil attributes, including humus content, available phosphorus and potassium, pH, particle size fractions, and exchangeable ion composition. Ensemble methods—particularly Gradient Boosting and Random Forest—were chosen due to their demonstrated ability to capture nonlinear relationships and variable interactions in soil systems. These models were evaluated using RMSE, MAE, MedianAE, and R^2 metrics, following standard machine-learning validation protocols. Geostatistical interpolation, including ordinary kriging and cokriging (Minasny and McBratney, 2016), was used to generate continuous surfaces and quantify spatial uncertainty. This combined machine-learning–geostatistics strategy enhanced predictive accuracy, allowed the reconstruction of sparsely sampled parameters, and improved model interpretability.

A substantial methodological component involved constructing semantic and ontological layers to represent soil knowledge within the digital twin. Drawing upon contemporary digital twin theory in engineering and CPSS frameworks (Tao et al., 2022), the system implemented a multi-dimensional semantic architecture that linked soil indicators, environmental drivers, agronomic practices, and system behaviors. Vector–graph knowledge representations were used to encode complex relationships among soil entities, enabling automated reasoning, cross-domain inference, and consistency checking. This semantic layer ensured that the digital twin operated not only as a data repository but as a knowledge-based system capable of interpreting environmental processes.

The computational infrastructure of the digital twin was built using a distributed and federated architecture. Edge computing nodes were utilized to preprocess IoT sensor data, reducing network load and ensuring low-latency updates. Federated learning frameworks—following concepts outlined in the Flower project—enabled the training of machine-learning models across decentralized data sources without transferring raw

datasets, ensuring data privacy and compliance with institutional requirements. Blockchain mechanisms, as described by Kalyaev and Melnik (2021) and Melnik et al. (2023b), were implemented to ensure transparent, immutable logging of data provenance, model updates, and system interactions. The integration of all data-ingestion pipelines, semantic layers, analytical modules, visualization tools, and computational processes into a structured, multi-tier architecture is illustrated in Figure 2, which represents the conceptual backbone of the digital twin.

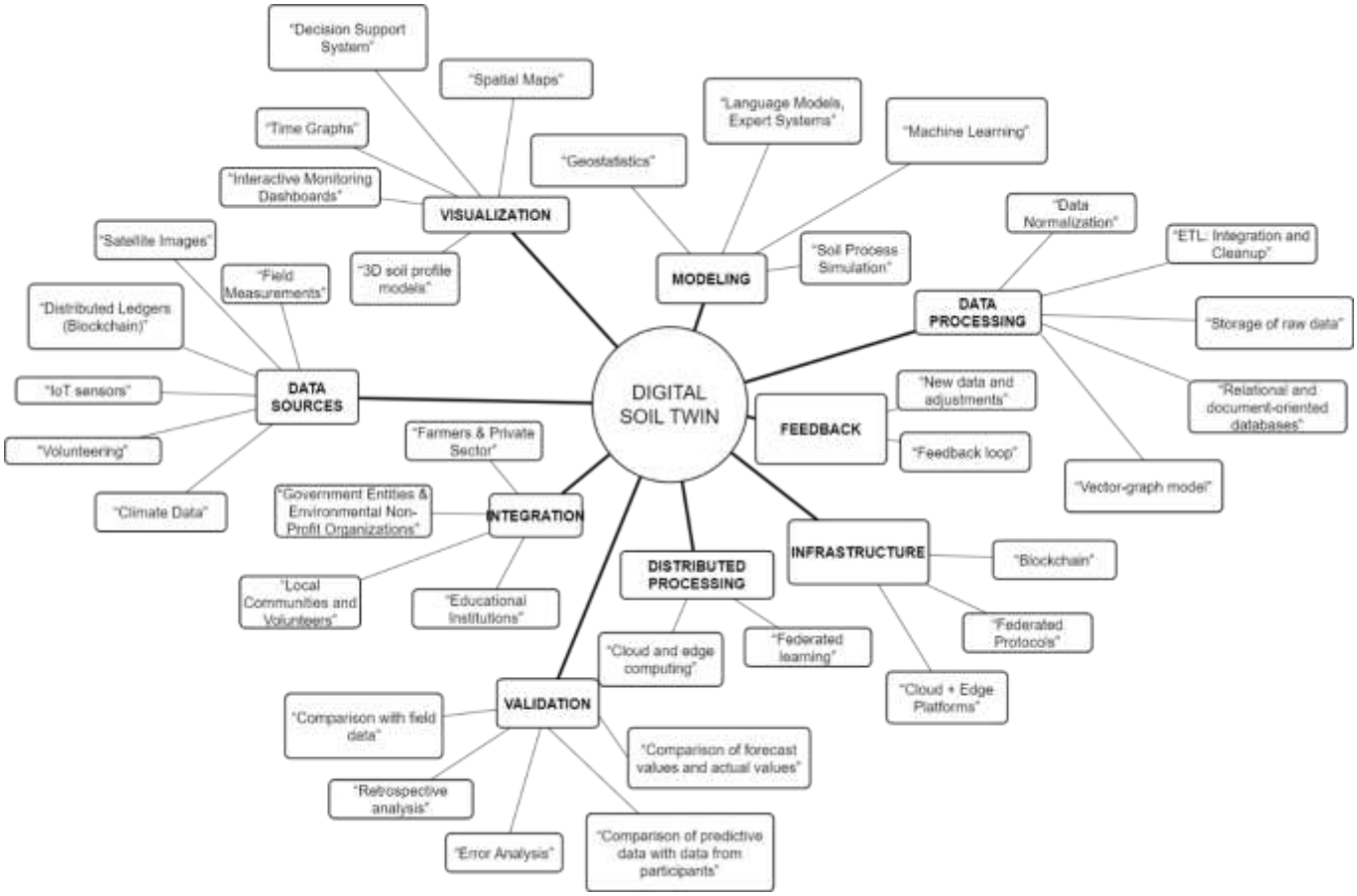


Figure 2. Conceptual architecture of the digital soil twin, including integrated data sources, distributed processing infrastructure, analytical and modeling modules, semantic knowledge structures, visualization components, validation mechanisms, and multi-directional feedback loops enabling continuous system refinement.

Visualization tools were developed to support multi-stakeholder engagement, including spatial distribution maps, 3D soil profile models, temporal dashboards, scenario simulation interfaces, and user-configurable decision-support tools. These components ensured that the digital twin could be actively utilized by researchers, agronomists, government agencies, private companies, and farmers, reflecting the model's design as a participatory socio-technical platform rather than a purely scientific simulation.

Results and Discussion

The systematic digitization, harmonization, and computational transformation of archival soil–geographical materials resulted in the construction of a unified, multi-layered soil database that integrates more than eight decades of observations across the Rostov region. This database, consisting of over one thousand vectorized soil polygons enriched with detailed morphological, analytical, and semantic descriptors, revealed substantial spatial heterogeneity in soil properties and enabled the reconstruction of long-term trends in soil fertility, anthropogenic impact, and land-use evolution. The integration of legacy datasets with contemporary observations underscored the value of historical materials while simultaneously highlighting the challenges associated with their analog origins, including inconsistent soil legends, varied coordinate reference systems, incomplete metadata, and gaps in analytical indicators such as exchangeable sodium, cation-exchange capacity, and micronutrient composition.

The machine-learning experiments conducted within the digital twin framework demonstrated that ensemble methods—particularly Gradient Boosting and Random Forest—provided consistently high predictive accuracy across most soil parameters. These findings are consistent with contemporary research on soil DT implementations, where ensemble models have been shown to effectively capture nonlinear

process interactions, high-dimensional relationships, and spatial context (Borovoy et al., 2024; Safonova et al., 2023). In the present study, parameters such as humus content, pH, available phosphorus, and available potassium were predicted with comparatively low RMSE and MAE values, reflecting both the abundance of historical measurements and the strong correlation of these indicators with environmental and land-use variables. Conversely, soil attributes for which historical measurements were sparse or irregular—such as exchangeable sodium, Ca^{2+}/Mg^{2+} ratios, and cation-exchange capacity—exhibited lower predictive performance, confirming earlier observations in the literature that data sparsity remains a structural constraint for soil-based machine-learning models (Silva et al., 2023; Chen et al., 2023).

Geostatistical modeling significantly improved the spatial resolution and interpretability of predicted soil attributes. Kriging and cokriging produced continuous surfaces with quantifiable uncertainty estimates, enabling the identification of zones susceptible to erosion, salinization, nutrient depletion, or structural degradation. These interpolated surfaces also allowed for the integration of predictive modeling with landscape-scale processes, supporting scenario analysis and risk assessment under varying climate conditions. The combined machine-learning–geostatistical approach aligns with best practices in digital soil mapping, where hybrid modeling has been widely recommended to reduce uncertainty and enhance spatial coherence (Minasny and McBratney, 2016).

The incorporation of remote sensing and IoT-based measurements into the digital twin enhanced the system's temporal sensitivity and allowed dynamic updates to soil state representations. Satellite indices reflecting vegetation vigor, soil brightness, and land surface moisture strengthened machine-learning predictions by supplying continuous spatial information in areas lacking ground measurements. Likewise, IoT-based soil moisture and temperature observations served as real-time inputs for model recalibration. These findings support conclusions from recent work demonstrating that integrating multi-source environmental data—particularly high-frequency sensor data—with DT architectures improves operational accuracy and predictive capacity (Nasirahmadi and Hensel, 2022; Tsakiridis et al., 2023).

A major methodological advance of the digital soil twin lies in its semantic and ontological architecture. By constructing a unified soil ontology and encoding soil–environment–management relationships using vector–graph knowledge representations, the DT was able to interpret soil indicators not as isolated measurements but as elements within a structured conceptual system. This semantic layer enabled automated reasoning functions, improved cross-domain interoperability, and facilitated the discovery of latent relationships within the data. These capabilities correspond with contemporary DT theory, which emphasizes the necessity of semantic integration, multi-source state fusion, and hierarchical knowledge modeling to support high-level decision-making in complex cyber–physical–social systems (Tao et al., 2022; Tagarakis et al., 2024). The successful application of these concepts to soil systems demonstrates the viability of transferring advanced engineering DT frameworks into agricultural and environmental domains.

The distributed computational architecture implemented in this study further contributed to the digital twin's performance and scalability. Edge computing nodes reduced latency by locally preprocessing IoT sensor data, while federated learning enabled model training across distributed datasets without requiring raw data transfer. This approach mitigated privacy concerns, reduced communication costs, and supported continuous model updates. The use of blockchain-based data-provenance mechanisms ensured transparent, immutable logging of data ingestion, model modifications, and system interactions, aligning with emerging best practices in critical-infrastructure DTs in sectors such as energy, transport, and manufacturing (Kalyaev and Melnik, 2021; Melnik et al. 2023a). The success of these implementations demonstrates that soil digital twins can benefit substantially from the same computational strategies used in industrial DT ecosystems.

The integration of all system components—archival datasets, contemporary measurements, machine-learning predictions, geostatistical surfaces, semantic knowledge graphs, and distributed computational tools—resulted in a functional virtual ecosystem capable of supporting real-time monitoring, scenario forecasting, and multi-stakeholder decision-making. Visualization tools, including 2D and 3D soil maps, temporal dashboards, simulation interfaces, and interactive profile viewers, made the system accessible to a wide range of users. As noted previously (see Figure 1), researchers, government agencies, private industries, agronomists, farmers, and citizens all represent target users of the soil digital twin. The present findings confirm that such a platform can simultaneously serve scientific, agronomic, regulatory, and socio-economic objectives. For instance, agronomists can identify nutrient-deficient fields; government agencies can detect degradation hotspots; private firms can virtually test fertilizer or irrigation treatments; and farmers can access real-time recommendations for irrigation scheduling, fertilization, and crop selection.

The results of this study therefore demonstrate the feasibility and scientific value of constructing a comprehensive soil digital twin for Southern Russia. By unifying historical archives with modern

observational technologies and advanced computational methods, the digital twin provides an analytically rich, continuously evolving, and operationally flexible representation of soil systems. The integration of heterogeneous data sources, multi-level semantic reasoning, hybrid ML–geostatistical modeling, and distributed processing places this work at the forefront of international soil DT research and positions the Rostov region as a pioneering case in large-scale agricultural digitalization.

Conclusion

The present study establishes a comprehensive theoretical, methodological, and computational foundation for developing a fully operational soil digital twin (DT) for Southern Russia, grounded in more than eighty years of soil–geographical survey archives, long-term agrochemical monitoring datasets, multi-scale soil maps, remote sensing products, IoT-based sensor networks, and advanced machine-learning and geostatistical tools. By integrating historically heterogeneous soil datasets with state-of-the-art digital technologies, the study demonstrates that the digital twin paradigm provides a transformative framework for real-time monitoring, scenario simulation, decision support, and long-term sustainability planning in agroecosystems. The results confirm that ensemble-based machine-learning methods—particularly Gradient Boosting and Random Forest—are highly effective for predicting numerous soil parameters, especially those with well-represented historical measurements. Their performance superiority reflects the inherently nonlinear nature of soil processes and the need for algorithms capable of modeling complex interactions between physical, chemical, and environmental attributes. Nevertheless, parameters characterized by sparse or inconsistent historical data, such as exchangeable sodium and cation-exchange capacity, remain challenging to model with high precision. These variables highlight the indispensable role of continuous data ingestion through IoT sensors, targeted sampling, and citizen-science platforms in strengthening the predictive capability of the digital soil twin.

The integration of geostatistics with machine learning significantly enhances the spatial continuity and uncertainty characterization of soil predictions. Techniques such as kriging and cokriging provide meaningful uncertainty quantification and spatial coherence, enabling robust assessment of erosion risks, nutrient deficiencies, hydrological stress, and anthropogenic degradation. At a higher conceptual level, the incorporation of semantic modeling, ontology-driven classification, and vector–graph knowledge structures advances the interoperability, interpretability, and inferential capacity of the digital twin, allowing explicit representation of relationships linking soil properties, climatic drivers, land-use practices, microbial processes, and ecological feedbacks. These semantic layers, enriched by insights from the reviewed literature, provide the structural backbone for a DT system capable of supporting cross-domain reasoning and automated decision support. The distributed computational architecture—encompassing edge computing, cloud–edge hybrid infrastructures, federated learning, and blockchain-based provenance systems—ensures that the digital twin can process large-scale data streams efficiently while maintaining data integrity, privacy, and transparency. Federated learning, in particular, allows geographically distributed sources to contribute to model training without exposing raw data, addressing one of the key challenges in agricultural digitalization. Blockchain-enabled immutability further ensures trust among heterogeneous stakeholders, including scientists, government agencies, farmers, and private-sector actors. These computational advances parallel global achievements in industrial and ecological digital twins and demonstrate their direct applicability to soil systems.

The digital soil twin developed through these principles has substantial practical value. It enables predictive modeling of crop responses under various irrigation, fertilization, and land-management scenarios, offering farmers real-time recommendations that can improve resource efficiency and increase yields by 15–25%, as shown in similar DT applications (Borovoy et al., 2024). It provides policymakers with reliable tools for assessing land degradation risks, guiding subsidy schemes, establishing sustainable farming programs, and improving land taxation frameworks based on spatially explicit fertility assessments. It supports environmental organizations in detecting early warning signals of erosion, salinization, or contamination. It enables insurance agencies to quantify risk and estimate expected losses using scenario-based stress tests. Furthermore, it strengthens scientific research by allowing soil scientists to explore long-term ecological trajectories, evaluate anthropogenic impacts, and validate process-based hypotheses through high-resolution spatiotemporal simulations.

Ultimately, the theoretical and methodological framework presented in this study positions the digital twin not simply as a technological tool but as a new epistemological model for soil science—one that unifies archival knowledge, real-time data, predictive modeling, semantic reasoning, and distributed intelligence. The integration of these components offers a powerful platform for advancing precision agriculture,

environmental monitoring, ecosystem resilience, and sustainable land-use planning across Southern Russia. As data density increases and computational capabilities expand, the digital soil twin will progressively evolve into a self-improving, adaptive system capable of supporting multi-sector decision-making, reducing uncertainty, and enhancing long-term agricultural and environmental sustainability.

Acknowledgements

The research was supported by the Strategic Academic Leadership Program of the Southern Federal University ("Priority 2030") and with the financial support of the Ministry of Science and Higher Education of Russia (Agreement No. 075-15-2025-667) using the equipment of the Soil Bioengineering Center for Collective use.

References

Bezuglova, O.S., 2009. Classification of soils. Publishing House of the Southern Federal University, Rostov-on-Don. 128p.

Bezuglova, O.S., Golozubov, O.M., Kryschenko, V.S., 2013. Soil-geographical large-scale electronic atlas of the Rostov region: principles of construction, structure, possibilities of use. SFedU, Rostov-on-Don. 146 p.

Borovoy, S.E., Komarova, O.P., Kozenko, K.Yu., 2024. Concept of digital twin of irrigated agrocenosis. *Izvestia of the Lower Volga Agro-University Complex* 3(75): 165–174.

Chen, G., Kang, X., Lin, M., Teng, S., Liu, Z., 2023. Stability prediction of soil slopes based on digital twinning and deep learning. *Applied Sciences* 13(11): 6470.

Derzhavina, L.M., Bulgakova, D.S., 2003. Methodical guidelines for conducting comprehensive monitoring of soil fertility of agricultural lands. FSBI "Rosinformagroteh", Moscow. 240 p.

ESA, 2023. The European Space Agency. Available at [Access date: 20.03.2025]: https://www.esa.int/

Eurostat, 2023. Overview: Land cover and use. Available at [Access date: 20.03.2025]: https://ec.europa.eu/eurostat/web/lucas

FSBI SCAS, 2025. Federal State Budgetary Institution State Center of Agrochemical Service. "Rostovsky". Available at [Access date: 08.01.2025]: http://www.donplodorodie.ru [in Russian].

Golozubov, O.M., 2013. Principles of creation of the soil-geographical electronic atlas of the Rostov region as a multifunctional reference and analytical system: Dissertation … candidate of biological sciences. Moscow: Lomonosov Moscow State University. 150 p.

Golozubov, O.M., Litvinov, Y.A., Kolesnikova, V.M., 2020. Methodical manual "Vectorization of large-scale soil maps". Electronic version. 72 p.

Gordov, E., 2023. Digital twins of systems and processes as a tool for modern climatology. *Fundamental and Applied Climatology* 9(3): 269–297.

Kalyaev, I.A., Melnik, E.V., 2021. Trusted control systems. Mekhatronika, Avtomatizatsiya, Upravlenie 22(5): 227–236.

Litvinov, Yu.A., 2018. Inventory, harmonization and analysis of heterogeneous soil–geographical data for the purposes of agroecological monitoring (on the example of the Rostov region): Dissertation … candidate of biological sciences. Moscow: Lomonosov Moscow State University. 145 p.

Melnik, E.V., Safronenkova, I.B., Taranov, A.Yu., 2023b. Ontological approach to solving the workload relocation problem in a distributed monitoring system with mobile components based on a distributed ledger. Izvestiya SFedU. *Engineering Sciences* 5(235): 163–173.

Melnik, E.V., Semenisty, S.A., Pukha, I.S., Gorelova, G.V., 2023a. Component of the analysis of the production environment parameters and trends in the production processes of the oil and gas industry. Certificate of registration of the computer program RU 2023613712, 17.02.2023.

Minasny, B., McBratney, A.B., 2016. Digital soil mapping: A brief history and some lessons. *Geoderma* 264(B): 301–311.

Nasirahmadi, A., Hensel, O., 2022. Toward the next generation of digitalization in agriculture based on digital twin paradigm. *Sensors* 22(2): 498.

Parewai, I., Köppen, M.A., 2025. Digital twin approach for soil moisture measurement with physically based rendering simulations and machine learning. *Electronics* 14(2): 395.

Petroselli, C., Williams, K.A., Ruiz, S.A., McKay Fletcher, D., Cooper, M.J., Roose, T., 2024. Microdialysis probes and digital twins reveal the rapid removal of fertiliser phosphate from the soil solution with an impact on crop nutrition in the short term. *Soil Biology and Biochemistry* 194: 109417.

Safonova, T.V., Kolbina, O.N., Yagotintseva, N.V., Mokryak, A.V., Istomin, Y.P., 2023. Application of the mathematical model of moisture loss by crop and soil in digital doubles (DD) of agroindustrial objects. *International Research Journal* 11: 137.

Silva, L., Rodríguez-Sedano, F., Baptista, P., Coelho, J.P., 2023. The digital twin paradigm applied to soil quality assessment: A systematic literature review. *Sensors* 23(7): 1007.

SoilWatch, 2022. Natural Climate Solution. Available at [Access date: 20.03.2025]: https://soilwatch.eu/

Tagarakis, A.C., Benos, L., Kyriakarakos, G., Pearson, S., Sørensen, C.G., Bochtis, D., 2024. Digital twins in agriculture and forestry: A review. *Sensors* 24(10): 3117.

Tao, F., Xiao, B., Qi, Q., Cheng, J., Ji, P., 2022. Digital twin modeling. *Journal of Manufacturing Systems* 64: 372–389.

Tsakiridis, N.L., Samarinas, N., Kalopesa, E., Zalidis, G.C., 2023. Cognitive soil digital twin for monitoring the soil ecosystem: A conceptual framework. *Soil Systems* 7(4): 88.