

Mechanistic Pedotransfer Functions for Soil Property Prediction Integrating Process Based Modeling and Data Driven Approaches

Caleb Ayomide Babatunde

babatundeayomide04@gmail.com

Ekiti State University

Research Article

Keywords: Mechanistic pedotransfer functions, Soil property prediction, Process-based modeling, Soil hydraulics, Precision agriculture

Posted Date: July 30th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10147346/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Pedotransfer functions (PTFs) have become indispensable tools in soil science for estimating difficult-to-measure soil properties from readily available data. Traditional PTFs have largely relied on empirical and statistical relationships, which often exhibit limited transferability beyond the conditions under which they were developed. In response, mechanistic pedotransfer functions (MPTFs) have emerged as a promising paradigm that integrates process-based understanding of soil physical, chemical, and biological processes with predictive modeling. This review synthesizes current knowledge on the conceptual foundations, theoretical frameworks, data requirements, modeling approaches, applications, and future prospects of mechanistic pedotransfer functions. Particular emphasis is placed on the representation of soil structure, pore-network dynamics, water flow, solute transport, carbon cycling, and root–soil interactions as fundamental drivers of soil property prediction. The review further examines process-based models, hybrid mechanistic–machine learning approaches, physics-informed artificial intelligence, digital soil twins, and emerging sensing technologies that are reshaping predictive soil science. Applications of mechanistic pedotransfer functions in hydraulic property estimation, irrigation management, land degradation assessment, soil carbon modeling, climate change studies, precision agriculture, and digital soil mapping are discussed. Additionally, major challenges related to data availability, scaling, uncertainty quantification, computational demands, and model transferability are critically evaluated. The synthesis demonstrates that mechanistic pedotransfer functions provide greater interpretability, physical realism, and extrapolation capability than conventional approaches, while recent advances in artificial intelligence and digital technologies offer new opportunities for enhanced prediction. Future developments are expected to focus on integrated mechanistic–AI frameworks, multi-source data fusion, and real-time soil monitoring systems that support sustainable land management and global food security.

1. Introduction

The accurate characterization of soil properties is fundamental to understanding soil functioning and supporting decision-making in agriculture, hydrology, environmental management, and climate change adaptation. Soil properties such as water retention characteristics, hydraulic conductivity, bulk density, organic carbon content, and cation exchange capacity govern key ecosystem processes including water movement, nutrient cycling, carbon sequestration, and plant growth (1, 2). However, direct measurement of these properties through laboratory analyses and field investigations is often laborious, expensive, and time-consuming, particularly when large spatial extents or high-resolution datasets are required. Consequently, the need for efficient approaches to estimate difficult-to-measure soil properties from readily available soil information has stimulated the development of pedotransfer functions (PTFs).

Pedotransfer functions, first formally conceptualized by (3), are predictive relationships that translate basic soil data, such as particle-size distribution, bulk density, and organic matter content, into estimates of more complex soil properties. Since their introduction, PTFs have become indispensable tools in soil science, enabling the parameterization of hydrological models, digital soil mapping frameworks, land

evaluation systems, and precision agriculture applications (4, 5). Traditional PTFs have largely relied on empirical or statistical relationships derived from regression analyses using observational datasets. While these approaches have demonstrated considerable utility, they often suffer from limited transferability across regions, inadequate representation of underlying soil processes, and poor performance outside the environmental conditions under which they were developed (6, 7).

Recent advances in soil physics, computational sciences, and environmental modeling have highlighted the need for next-generation PTFs that transcend purely empirical correlations. The inherent complexity of soil systems arises from interactions among soil texture, structure, mineralogy, biological activity, and environmental drivers operating across multiple spatial and temporal scales (8). Empirical models frequently fail to capture these mechanistic relationships, limiting their predictive robustness under changing climatic conditions and novel management scenarios. As a result, increasing attention has shifted toward mechanistic pedotransfer functions that incorporate process-based understanding of soil behavior into predictive frameworks.

Mechanistic pedotransfer functions seek to estimate soil properties by integrating physically based principles governing soil processes, including pore-scale water flow, capillary dynamics, soil structural organization, and biophysical interactions. Unlike conventional PTFs that depend solely on statistical associations, mechanistic approaches leverage knowledge of the causal mechanisms controlling soil functions, thereby enhancing model interpretability and extrapolation capability (9, 10). Such frameworks often employ physically based equations, pore-network models, and process-based hydrological models to establish relationships between measurable soil attributes and target soil functions.

The emergence of advanced computational techniques has further expanded the scope of mechanistic pedotransfer modeling. Hybrid approaches combining mechanistic principles with machine learning algorithms, physics-informed neural networks, and data assimilation techniques have shown considerable promise in improving prediction accuracy while preserving physical realism (11, 12). These developments align closely with the broader digital transformation occurring within soil science, characterized by increasing adoption of digital soil mapping, proximal sensing technologies, high-resolution imaging, and artificial intelligence-based decision support systems (13, 14).

Mechanistic PTFs are particularly relevant in addressing contemporary challenges associated with sustainable land management and climate resilience. Accurate predictions of soil hydraulic properties are essential for optimizing irrigation strategies, evaluating drought vulnerability, assessing flood risks, and modeling ecosystem responses to climate change (15). Similarly, improved understanding of soil physical and biological processes contributes to enhanced assessments of carbon dynamics, nutrient cycling, and soil health indicators that underpin sustainable agricultural intensification and ecosystem conservation.

Despite their growing importance, mechanistic pedotransfer functions remain an emerging field characterized by methodological diversity, data limitations, computational challenges, and unresolved questions concerning model generalizability and scalability. Existing reviews have predominantly focused

on empirical PTF development, machine learning applications, or soil hydraulic modeling independently, with limited synthesis of the conceptual and methodological integration between process-based understanding and modern data-driven approaches. Consequently, there is a pressing need for a comprehensive review that critically examines the evolution, theoretical foundations, modeling frameworks, applications, challenges, and future directions of mechanistic pedotransfer functions within contemporary soil science.

Therefore, this review aims to bridge the gap between traditional pedotransfer methodologies and emerging mechanistic paradigms by synthesizing current knowledge on process-based soil property prediction. Specifically, the review seeks to: (i) examine the conceptual evolution of pedotransfer functions from empirical to mechanistic frameworks; (ii) evaluate the soil physical and biological principles underpinning mechanistic PTF development; (iii) explore the integration of process-based models with data-driven techniques, including machine learning and artificial intelligence; (iv) assess current applications of mechanistic PTFs across soil and environmental disciplines; and (v) identify existing challenges, research gaps, and future opportunities for advancing predictive soil science. Through this synthesis, the review aims to provide a foundation for developing robust, interpretable, and transferable pedotransfer frameworks capable of supporting sustainable soil management in an era of rapid environmental change.

2. Methodology

This review was conducted through a comprehensive examination of published literature on pedotransfer functions, process-based soil modeling, soil physics, hydrology, digital soil mapping, and artificial intelligence applications in soil science. Relevant peer-reviewed journal articles, books, conference proceedings, and technical reports published between 1976 and 2026 were sourced from major scientific databases including Google Scholar, Scopus, Web of Science, ScienceDirect, and SpringerLink using keywords such as *mechanistic pedotransfer functions*, *soil hydraulic properties*, *process-based modeling*, *physics-informed machine learning*, *digital soil twins*, and *soil prediction*.

The collected literature was critically evaluated and synthesized to examine the conceptual foundations, modeling frameworks, data requirements, applications, technological advances, uncertainties, and future directions of mechanistic pedotransfer functions. Emphasis was placed on studies that integrated soil process understanding with predictive modeling approaches, thereby providing insights into the evolution of pedotransfer science from empirical relationships toward mechanistic and hybrid predictive frameworks.

3. Conceptual Foundations of Pedotransfer Functions

3.1 Definition and Principles of Pedotransfer Functions

Pedotransfer functions (PTFs) are predictive relationships that estimate difficult-to-measure soil properties from more readily available soil information. (3) originally defined PTFs as tools that "translate

data we have into what we need," emphasizing their utility in converting routinely measured soil properties into parameters required for land evaluation and process-based modeling. Since then, PTFs have become integral components of soil hydrology, digital soil mapping, environmental modeling, and precision agriculture (4, 5).

The underlying principle of PTFs is that many soil properties are intrinsically linked through pedogenic processes and physical interactions. Readily measured properties such as particle-size distribution, bulk density, organic matter content, and soil structure can therefore be used to infer hydraulic, chemical, and biological characteristics that are more difficult or expensive to determine directly (6). For example, soil texture influences pore-size distribution, which subsequently controls water retention and hydraulic conductivity (16). Similarly, organic matter contributes to aggregate stability and porosity, thereby affecting soil hydraulic behavior (17).

Several principles guide the development of PTFs. First, predictive relationships should be based on robust associations between predictor and target variables. Second, the quality and representativeness of calibration datasets strongly influence model reliability and transferability (9). Third, model validation using independent datasets is essential to assess predictive performance. Finally, modern PTF development increasingly emphasizes incorporating physical understanding of soil processes alongside statistical predictive capability (5).

3.2 Historical Development of Pedotransfer Modeling

The evolution of pedotransfer modeling reflects advances in soil science, statistical methodologies, and computational technologies. Before the formal introduction of PTFs, soil scientists often relied on expert judgment and qualitative interpretations derived from soil survey information to infer soil behavior. Although useful, these approaches lacked consistency and quantitative rigor.

The formal development of PTFs began with (3), who proposed using readily available soil survey data to estimate parameters required for simulation models. Early PTFs focused primarily on soil hydraulic properties, particularly water retention characteristics and saturated hydraulic conductivity. Pioneering studies by (18–20) developed regression-based equations relating soil texture and organic matter content to hydraulic parameters.

During the 1990s, increasing availability of soil databases facilitated the application of multivariate statistical techniques, including multiple linear regression and discriminant analysis, to improve prediction accuracy (21). The publication of HYPRES, a European database of soil hydraulic properties, represented a major milestone by providing standardized datasets for developing and validating PTFs across diverse environmental conditions (21).

The early 2000s witnessed a conceptual expansion from conventional PTFs toward soil inference systems (SINFERS), which integrated multiple predictive models to estimate various soil functions and behaviors (4). Concurrently, advances in computational power facilitated the adoption of machine

learning techniques, including artificial neural networks and decision tree algorithms, capable of modeling complex nonlinear relationships.

Recent developments have focused increasingly on mechanistic and hybrid approaches that combine physical understanding of soil processes with data-driven methods. Emerging frameworks incorporating physics-informed machine learning, digital soil mapping, and artificial intelligence seek to improve predictive accuracy while maintaining interpretability and physical realism (12, 14). Thus, the historical trajectory of PTF development reflects a progression from empirical observations toward integrated predictive systems grounded in both data and process understanding.

3.3 Classification of Pedotransfer Functions

Pedotransfer functions can be classified according to the methodological approaches used in their development. Four major categories are commonly recognized: empirical PTFs, statistical PTFs, machine learning-based PTFs, and mechanistic PTFs. These categories represent an evolutionary continuum from simple observational relationships to sophisticated process-based prediction frameworks.

3.3.1 Empirical Pedotransfer Functions

Empirical PTFs constitute the earliest generation of pedotransfer models and are based primarily on direct observational relationships derived from experimental datasets. These functions commonly employ simple regression equations to predict target soil properties from readily measurable variables such as texture, bulk density, and organic matter content (6).

The simplicity of empirical PTFs makes them attractive for practical applications. They require relatively limited datasets, are computationally efficient, and offer straightforward interpretation. Widely used examples include the equations developed by (16, 20) for estimating soil water retention characteristics.

Despite their utility, empirical PTFs possess important limitations. Their applicability is often restricted to environmental conditions similar to those represented in the calibration datasets. Moreover, because these models rely solely on observed correlations, they may not adequately capture the underlying physical mechanisms governing soil behavior (5).

3.3.2 Statistical Pedotransfer Functions

Statistical PTFs extend empirical approaches by employing more advanced statistical techniques to improve predictive performance and uncertainty characterization. Common methods include multiple linear regression, generalized linear models, principal component analysis, discriminant analysis, and Bayesian inference (1).

These approaches enable simultaneous consideration of multiple predictor variables and interactions among soil attributes. Bayesian frameworks are particularly advantageous because they allow incorporation of prior knowledge while quantifying uncertainty associated with predictions (22).

However, statistical PTFs remain fundamentally dependent on correlations derived from calibration datasets. Consequently, their predictive performance may decline when applied to soils or environmental conditions substantially different from those used during model development. Furthermore, statistical models generally provide limited insight into the causal mechanisms driving soil processes.

3.3.3 Machine Learning-Based Pedotransfer Functions

Machine learning-based PTFs utilize computational algorithms capable of identifying complex nonlinear relationships within large datasets without requiring predefined mathematical forms. Techniques frequently employed include artificial neural networks, support vector machines, random forests, gradient boosting algorithms, k-nearest neighbors, and deep learning models (14).

Machine learning approaches have demonstrated superior predictive performance compared with conventional regression methods in many soil science applications (7). Their ability to integrate diverse datasets, including proximal sensing measurements, remote sensing products, terrain attributes, and climatic variables, enhances their applicability within digital soil mapping frameworks.

Nevertheless, machine learning models often function as "black-box" systems with limited interpretability. Their reliance on large, representative datasets and susceptibility to overfitting present additional challenges. Moreover, predictions generated outside the range of training conditions may be unreliable, raising concerns regarding model extrapolation and transferability (11).

3.3.4 Mechanistic Pedotransfer Functions

Mechanistic pedotransfer functions represent an emerging paradigm emphasizing physical understanding of soil processes as the basis for prediction. Unlike empirical and machine learning approaches that rely predominantly on statistical associations, mechanistic PTFs seek to incorporate causal relationships governing soil behavior, including pore-scale flow processes, capillary phenomena, and structural organization (9).

These approaches frequently employ physically based theories such as capillary bundle models, pore-network analyses, and hydrological flow equations to estimate hydraulic properties from measurable soil characteristics (10). Advances in soil imaging technologies, including X-ray computed tomography, have further enhanced opportunities to integrate pore architecture directly into predictive frameworks.

Mechanistic PTFs offer several advantages, including improved interpretability, stronger theoretical foundations, and enhanced capacity for extrapolation under changing environmental conditions. However, they often require more detailed input data and greater computational resources than conventional approaches. Their development remains an active research frontier aimed at balancing physical realism with operational feasibility (5).

3.4 Theoretical Basis for Mechanistic Soil Property Prediction

Mechanistic soil property prediction is founded on the premise that soil functions arise from interactions among physical, chemical, and biological processes operating across multiple spatial and temporal scales. Unlike empirical approaches that focus on statistical associations, mechanistic frameworks aim to represent the processes responsible for generating observed soil behaviors (8).

A fundamental principle underlying mechanistic prediction is the relationship between soil structure and function. Soil hydraulic properties are controlled by pore-size distribution, connectivity, tortuosity, and aggregate arrangement rather than solely by texture composition (23). Consequently, mechanistic PTFs increasingly emphasize characterization of soil pore architecture as a basis for predicting water flow and retention behavior.

Several theoretical concepts support mechanistic prediction. Capillary theory describes water retention as a function of pore radius distribution, whereas Darcy's law provides a framework for understanding saturated flow through porous media. Richards' equation integrates capillary and gravitational forces to describe unsaturated water movement within soils (24). Additionally, pore-network models explicitly represent three-dimensional pore systems and their influence on hydraulic conductivity and solute transport.

Recent advances in imaging technologies and computational modeling have enabled direct observation and simulation of pore-scale processes, thereby strengthening the mechanistic basis of soil property prediction (23). Furthermore, recognition of dynamic interactions involving root growth, microbial activity, wetting–drying cycles, and land management practices has expanded mechanistic frameworks beyond purely physical considerations.

The emergence of hybrid approaches integrating mechanistic understanding with machine learning techniques represents a significant advancement in predictive soil science. Physics-informed machine learning models constrain predictions using established physical laws while exploiting the flexibility of data-driven algorithms (12). Such frameworks offer promising opportunities to improve prediction accuracy, interpretability, and transferability, thereby supporting sustainable soil management under increasingly variable environmental conditions.

4. Mechanistic Understanding of Soil Processes

4.1 Soil Formation Processes and Property Development

The mechanistic prediction of soil properties requires a fundamental understanding of soil genesis and the processes responsible for the development of soil characteristics. Soil formation is a dynamic and continuous process involving interactions among climate, organisms, relief, parent material, and time, collectively described by the CLORPT framework proposed by (25, 26). These factors influence the physical, chemical, and biological transformations that produce distinct soil horizons and determine soil functionality. Since pedotransfer functions aim to predict soil properties from readily measurable

attributes, understanding the processes responsible for generating those attributes provides the theoretical basis for mechanistic prediction.

Soil formation begins with the weathering of parent materials through physical, chemical, and biological processes. Physical weathering disintegrates rocks into smaller particles through temperature fluctuations, wetting and drying cycles, freeze–thaw processes, and mechanical abrasion. Chemical weathering alters mineral composition through hydrolysis, oxidation, reduction, dissolution, and carbonation reactions, producing secondary clay minerals and releasing nutrients into the soil system (27). Biological weathering further accelerates mineral decomposition through root activity, microbial metabolism, and organic acid production. Together, these processes determine the texture, mineralogy, and geochemical composition of developing soils.

Pedogenic processes subsequently modify weathered materials and contribute to horizon differentiation. Four fundamental pedogenic processes—additions, losses, transformations, and translocations—govern soil development (28). Additions include inputs of organic residues, dust, sediments, and dissolved substances, whereas losses involve erosion, leaching, and gaseous emissions. Transformations encompass organic matter decomposition, mineral weathering, and aggregate formation, while translocations refer to the movement of clay, organic compounds, carbonates, and soluble ions within the soil profile. These processes interact continuously to create the morphological and functional characteristics observed in soils.

The development of soil physical properties is closely linked to pedogenic evolution. Soil texture, determined primarily by the proportions of sand, silt, and clay, originates largely from parent material and weathering processes. However, soil structure develops through biological activity, organic matter accumulation, shrink–swell dynamics, and aggregation processes (29). Structural development strongly influences porosity, pore connectivity, hydraulic conductivity, and water retention, which are among the most frequently predicted properties in pedotransfer modeling.

Chemical properties also evolve through pedogenesis. Weathering reactions release primary nutrients such as potassium, calcium, magnesium, and phosphorus, while clay mineral formation influences cation exchange capacity and nutrient retention. Organic matter accumulation contributes significantly to nutrient cycling, buffering capacity, and soil fertility. In highly weathered tropical soils, prolonged leaching may result in nutrient depletion and dominance of iron and aluminum oxides, thereby influencing soil chemical behavior and plant nutrient availability (30).

Biological processes constitute another critical dimension of soil development. Microorganisms, plant roots, and soil fauna actively participate in decomposition, nutrient mineralization, aggregation, and bioturbation. These organisms influence both soil structure and chemical transformations, generating feedback mechanisms that regulate ecosystem functioning (8). Biological activity also contributes to the formation of biopores and aggregate-associated organic matter, which affect water flow pathways and carbon stabilization.

The influence of soil-forming factors extends across multiple spatial and temporal scales. Climatic variables such as precipitation and temperature regulate weathering intensity, organic matter decomposition, and leaching rates. Topography controls drainage patterns, erosion, deposition, and microclimatic conditions, leading to systematic variations in soil properties along toposequences. Vegetation influences organic matter inputs, rooting patterns, and biological activity, while parent material determines initial mineral composition and weathering potential. Time integrates the cumulative effects of these factors, producing increasingly differentiated soil profiles as pedogenesis progresses (25, 26).

Mechanistic pedotransfer functions seek to exploit these process-based relationships by linking measurable indicators of soil formation to functional soil properties. For example, texture reflects weathering history, organic carbon content reflects biological accumulation and decomposition processes, and horizon morphology reflects translocation mechanisms. By incorporating pedogenic understanding into predictive frameworks, mechanistic PTFs can improve both interpretability and transferability compared with purely empirical approaches. Consequently, soil formation processes provide the foundational context for understanding the origin, variability, and predictability of soil properties across landscapes.

4.2 Soil Structure and Pore Network Dynamics

Soil structure refers to the spatial arrangement of primary soil particles into aggregates and the organization of pore spaces between and within those aggregates. Unlike soil texture, which is relatively stable and inherited largely from parent material, soil structure is dynamic and continuously modified by biological activity, wetting–drying cycles, tillage operations, root growth, and organic matter turnover (31). Because hydraulic, aeration, and nutrient transport processes are strongly governed by pore architecture, soil structure occupies a central position in mechanistic soil property prediction.

The pore system of soils consists of a complex network of interconnected voids varying in size, shape, tortuosity, and connectivity. Macropores ($> 75 \mu\text{m}$) facilitate rapid water infiltration, preferential flow, and gaseous exchange, whereas mesopores and micropores primarily regulate water storage and retention (32). The distribution and connectivity of these pores determine the movement of water, solutes, gases, and biological organisms throughout the soil profile. Consequently, two soils with similar textures may exhibit markedly different hydraulic behavior because of differences in structural organization.

Recent advances in X-ray computed tomography (CT), magnetic resonance imaging, and image analysis have revealed that soil hydraulic properties emerge directly from three-dimensional pore architecture (23). Pore-network models have subsequently become valuable tools for simulating fluid flow and predicting hydraulic conductivity, water retention, and gas diffusion. These models explicitly represent soil pore geometry and connectivity, providing a mechanistic framework that extends beyond traditional texture-based predictions.

Aggregation processes further influence pore dynamics. Soil aggregates are formed through interactions among clay particles, organic matter, roots, fungal hyphae, and microbial exudates (29). Stable aggregates create hierarchical pore systems that improve water retention, enhance infiltration, and

reduce erosion susceptibility. The stability of these aggregates therefore plays a critical role in maintaining soil functionality and resilience.

Mechanistic pedotransfer functions increasingly incorporate indicators of soil structure, including aggregate stability, porosity, bulk density, and pore-size distribution, because these properties exert direct control over hydraulic behavior. By integrating structural descriptors into predictive frameworks, mechanistic models can more accurately represent the processes governing soil water dynamics than conventional empirical approaches based solely on texture.

4.3 Water Flow and Retention Mechanisms

Water movement and storage constitute fundamental soil functions that directly influence plant growth, nutrient availability, and ecosystem productivity. Consequently, prediction of soil hydraulic properties has historically been one of the principal applications of pedotransfer functions. Mechanistic approaches seek to represent the physical processes controlling water flow and retention rather than relying exclusively on statistical relationships.

Water retention in soils is governed primarily by capillary and adsorptive forces. Following saturation, gravitational drainage removes water from larger pores, while capillary forces retain water within smaller pores. The relationship between soil water content and matric potential is described by the soil water retention curve, which serves as a fundamental indicator of hydraulic behavior (33, 34). Soil texture, structure, organic matter content, and pore-size distribution collectively determine the shape of this curve.

Water flow through saturated soils is commonly described by Darcy's law, which relates water flux to hydraulic gradients and hydraulic conductivity. Under unsaturated conditions, water movement becomes considerably more complex because hydraulic conductivity decreases as soil moisture declines. Richards' equation integrates Darcy's law with continuity principles to describe unsaturated water flow and forms the basis of many process-based hydrological models (35).

Pore-scale heterogeneity significantly influences hydraulic behavior. Water does not move uniformly through soils but instead follows preferential pathways associated with macropores, root channels, cracks, and biological pores. These preferential flow mechanisms often result in rapid transport that cannot be adequately represented by conventional equilibrium models (36). Mechanistic pedotransfer functions therefore increasingly account for pore connectivity and structural heterogeneity.

Advances in soil physics have enabled the development of physically based hydraulic models linking pore-size distributions directly to water retention and conductivity characteristics. Such approaches provide a stronger theoretical basis for predicting hydraulic properties and improve extrapolation beyond calibration datasets.

4.4 Solute Transport Processes

The movement of dissolved substances through soil is governed by a combination of physical, chemical, and biological processes collectively referred to as solute transport. Understanding these mechanisms is essential for predicting nutrient availability, contaminant mobility, agrochemical behavior, and groundwater quality.

Solute transport occurs primarily through advection, diffusion, and dispersion. Advection refers to the movement of solutes with flowing water, while diffusion involves movement from regions of higher concentration to lower concentration. Mechanical dispersion results from variations in flow velocities within different pore pathways, causing spreading of solute plumes (37). Together, these processes determine the spatial and temporal distribution of dissolved substances within soils.

Chemical interactions significantly influence solute mobility. Adsorption-desorption reactions, ion exchange processes, precipitation-dissolution reactions, and complexation mechanisms regulate the retention and release of nutrients and contaminants. Clay minerals, organic matter, and oxide surfaces provide reactive sites capable of immobilizing or releasing dissolved species depending on environmental conditions (38, 39).

Biological processes further modify solute dynamics through nutrient uptake, mineralization, immobilization, and microbial transformation. Nitrogen cycling provides a prominent example in which microbial processes regulate the conversion of organic nitrogen into ammonium and nitrate, thereby influencing nutrient availability and leaching risks.

Mechanistic pedotransfer functions aimed at predicting solute transport increasingly integrate hydraulic properties with chemical and biological descriptors. Such approaches provide a more comprehensive representation of transport processes than conventional empirical models and are particularly valuable for environmental risk assessment and nutrient management applications.

4.5 Soil Organic Matter Dynamics and Carbon Cycling

Soil organic matter (SOM) is a critical determinant of soil quality and ecosystem functioning. It influences nutrient cycling, aggregation, water retention, microbial activity, and carbon sequestration. Mechanistic prediction of soil properties therefore requires an understanding of the processes governing SOM accumulation, transformation, and stabilization.

Soil organic matter originates primarily from plant residues, root exudates, microbial biomass, and animal remains. Following deposition, these materials undergo decomposition mediated by soil microorganisms. During decomposition, organic compounds are transformed into simpler molecules, releasing nutrients and carbon dioxide while contributing to the formation of humified organic matter (40).

The persistence of SOM depends on several stabilization mechanisms. Physical protection within aggregates, chemical association with mineral surfaces, and biochemical recalcitrance collectively regulate organic matter turnover rates. Recent studies suggest that mineral-organic associations play a more significant role in long-term carbon stabilization than previously recognized (41, 42).

Soil organic matter exerts substantial influence on physical properties. Increased SOM generally promotes aggregate formation, enhances porosity, reduces bulk density, and improves water-holding capacity (17). Consequently, organic carbon content is among the most frequently utilized predictors in pedotransfer modeling.

Carbon cycling also links soils to global climate systems. Soils contain approximately three times more carbon than the atmosphere and represent a major component of the terrestrial carbon cycle (43). Accurate prediction of carbon dynamics is therefore essential for climate mitigation strategies and sustainable land management. Mechanistic PTFs increasingly integrate carbon cycling processes to improve prediction of soil functional properties under changing environmental conditions.

4.6 Root–Soil Interactions and Rhizosphere Processes

The rhizosphere, defined as the narrow zone of soil influenced by plant roots, represents one of the most biologically active regions within terrestrial ecosystems. Root–soil interactions influence soil structure, hydraulic properties, nutrient availability, microbial activity, and carbon cycling, making them highly relevant to mechanistic soil property prediction.

Roots modify soil structure through physical penetration, pore creation, and aggregate stabilization. As roots grow, they exert mechanical forces that alter pore geometry and connectivity. Following root decay, biopores remain and often serve as preferential pathways for water infiltration and root exploration by subsequent crops (44). These structural modifications directly influence hydraulic conductivity and water retention characteristics.

Root exudates further affect soil processes by supplying carbon substrates that stimulate microbial activity. These compounds promote microbial growth, aggregation, and nutrient transformations within the rhizosphere. Enhanced microbial activity contributes to the formation of stable aggregates and influences nutrient availability through mineralization and immobilization processes (45).

Root water uptake constitutes another critical process influencing soil hydraulic behavior. Water extraction by roots generates moisture gradients around root surfaces, affecting local hydraulic conductivity and water distribution patterns. Modern rhizosphere models increasingly integrate root architecture, water uptake dynamics, and soil hydraulic properties to simulate plant–soil interactions more accurately.

Nutrient acquisition processes are also strongly influenced by rhizosphere dynamics. Roots alter soil pH, release enzymes, and interact with mycorrhizal fungi to enhance nutrient availability. These interactions contribute to the spatial heterogeneity of soil chemical properties and influence nutrient transport processes.

Because root activity continuously modifies soil physical, chemical, and biological characteristics, incorporation of rhizosphere processes represents a critical frontier in mechanistic pedotransfer

modeling. Integrating root–soil interactions into predictive frameworks offers opportunities to improve predictions of hydraulic behavior, nutrient cycling, and ecosystem responses to environmental change.

5. Mechanistic Pedotransfer Functions: Principles and Frameworks

5.1 Concept and Definition of Mechanistic Pedotransfer Functions

Mechanistic pedotransfer functions (MPTFs) represent an emerging class of predictive tools that estimate soil properties by explicitly incorporating the physical, chemical, and biological processes governing soil behavior. Unlike conventional pedotransfer functions, which primarily rely on empirical correlations or statistical associations between readily measured and difficult-to-measure soil properties, MPTFs seek to establish causal relationships rooted in soil process understanding (5, 9). Their development reflects a paradigm shift from purely data-driven prediction toward process-informed modeling capable of improving interpretability, robustness, and extrapolation potential.

The concept of mechanistic pedotransfer functions emerged in response to limitations associated with empirical and machine-learning-based PTFs. While empirical models often provide satisfactory predictions within the range of calibration data, their performance frequently deteriorates when applied to different climatic conditions, soil types, or management systems (6). Mechanistic approaches address this challenge by embedding soil physical laws and process understanding into predictive frameworks, thereby reducing dependence on site-specific correlations.

Mechanistic PTFs are founded on the premise that measurable soil attributes represent manifestations of underlying processes controlling soil functioning. For example, soil texture influences pore-size distribution, which subsequently governs water retention and hydraulic conductivity. Rather than directly relating texture to hydraulic conductivity through statistical regression, a mechanistic approach may first estimate pore-size distribution and then derive hydraulic properties using physical flow equations (46, 47). This intermediate representation enhances the scientific basis of prediction and facilitates model transferability.

The increasing availability of high-resolution soil datasets, imaging technologies, and computational resources has accelerated interest in mechanistic PTFs. These developments allow researchers to characterize pore architecture, soil structure, and dynamic processes with unprecedented detail, thereby enabling predictive models that more accurately reflect the complexity of soil systems (23). Consequently, MPTFs are increasingly viewed as a critical component of next-generation predictive soil science.

5.2 Physical Basis of Mechanistic Prediction

The physical basis of mechanistic pedotransfer modeling lies in the recognition that soil functions emerge from interactions among particles, pores, fluids, and biological agents operating across multiple

scales. These interactions are governed by well-established physical laws describing water flow, gas movement, heat transfer, and mechanical behavior in porous media (24, 33).

One of the most important physical foundations of mechanistic prediction is capillary theory. Soil pores behave as capillary tubes in which water retention depends on pore radius and surface tension. According to the Young–Laplace equation, smaller pores retain water more strongly than larger pores. Consequently, pore-size distribution serves as a direct determinant of soil water retention characteristics (48). Mechanistic PTFs exploit this relationship by linking measurable soil properties to pore geometry and subsequently to hydraulic behavior.

Darcy's law provides another cornerstone of mechanistic prediction. The law describes water flow through porous media as a function of hydraulic conductivity and hydraulic gradients. Under unsaturated conditions, Richards' equation extends Darcy's law by accounting for variations in soil moisture content and matric potential (35). Many mechanistic hydraulic models employ these equations to estimate water movement based on physically meaningful soil parameters.

Additionally, theories describing pore connectivity, tortuosity, and network topology contribute significantly to mechanistic prediction. Water and solutes do not move through soils in straight pathways but rather follow complex interconnected pore networks. The degree of connectivity among pores strongly influences infiltration rates, hydraulic conductivity, and transport behavior (23). Therefore, mechanistic PTFs increasingly incorporate pore-network concepts to improve predictive realism.

The physical basis of MPTFs also extends to thermal and gaseous transport processes. Soil temperature regimes, gas diffusion, and aeration dynamics are controlled by porosity, moisture content, and pore continuity. Mechanistic representations of these processes enhance the ability of predictive models to simulate soil functioning under varying environmental conditions.

5.3 Process-Based Representation of Soil Properties

A defining characteristic of mechanistic pedotransfer functions is their reliance on process-based representations of soil properties. Rather than treating soil properties as isolated variables, MPTFs conceptualize them as outcomes of interacting pedogenic, hydraulic, biological, and geochemical processes (8).

Water retention provides a useful example. Traditional PTFs often estimate field capacity or wilting point directly from texture and organic matter content. Mechanistic approaches, however, represent water retention as an emergent property resulting from pore-size distribution, pore connectivity, surface area, and capillary forces (9). By explicitly accounting for these mechanisms, predictions become more physically meaningful and transferable.

Similarly, saturated and unsaturated hydraulic conductivity are represented as functions of pore geometry, continuity, and flow pathways. Soil structure, aggregate arrangement, and biological pores influence the

ease with which water moves through soil profiles. Mechanistic frameworks therefore seek to characterize these controlling factors rather than relying solely on empirical relationships.

Chemical properties can also be represented mechanistically. Cation exchange capacity, nutrient retention, and contaminant sorption depend on mineralogy, organic matter composition, and surface charge characteristics (39, 49). Mechanistic models incorporate these interactions to improve prediction of chemical behavior under varying environmental conditions.

Biological processes represent another important dimension. Root growth, microbial activity, decomposition, and aggregation continuously modify soil properties over time. Contemporary MPTFs increasingly integrate biological dynamics, recognizing that soil functions cannot be fully understood through physical processes alone (40).

By representing soil properties as outcomes of interconnected processes, mechanistic PTFs provide a more comprehensive understanding of soil functioning and facilitate predictions under novel environmental scenarios.

5.4 Incorporation of Soil Structural Information

Soil structure constitutes one of the most critical determinants of soil behavior and is therefore central to mechanistic pedotransfer modeling. Structure refers to the arrangement of soil particles into aggregates and the organization of pores within the soil matrix. Because hydraulic, aeration, and transport properties depend heavily on pore architecture, incorporation of structural information significantly enhances predictive performance (31).

Traditional PTFs have often relied on texture as a surrogate for structure because structural measurements are more difficult to obtain. However, soils with similar textures can exhibit markedly different hydraulic behaviors due to differences in aggregation, porosity, and pore connectivity. Consequently, mechanistic PTFs increasingly utilize direct or indirect indicators of soil structure, including bulk density, aggregate stability, porosity, penetration resistance, and imaging-derived metrics.

Recent advances in X-ray computed tomography (CT) have revolutionized structural characterization by enabling visualization of three-dimensional pore networks (23). CT-derived metrics such as pore volume, connectivity, tortuosity, and fractal dimensions provide valuable information for mechanistic prediction of hydraulic properties. These datasets allow researchers to develop pore-network models that directly simulate water flow and retention processes.

Biological structures also contribute significantly to soil architecture. Root channels, earthworm burrows, and microbial aggregates create preferential flow pathways that strongly influence hydraulic conductivity and solute transport (50). Incorporating these biologically generated structures into predictive frameworks improves representation of real-world soil behavior.

The integration of structural information therefore represents a major advancement over conventional texture-based approaches and constitutes a key distinguishing feature of mechanistic pedotransfer

modeling.

5.5 Scaling Issues in Mechanistic Pedotransfer Modeling

Scaling remains one of the most significant challenges in mechanistic pedotransfer modeling. Soil processes occur across a wide range of spatial and temporal scales, from nanometer-scale mineral interactions to landscape-scale hydrological processes. Translating information across these scales is essential for developing predictive frameworks that are both physically realistic and operationally useful (51).

At the pore scale, hydraulic and transport processes are governed by pore geometry, surface interactions, and capillary forces. Advances in imaging technologies have enabled detailed characterization of these processes. However, measurements obtained at microscopic scales may not directly represent field-scale behavior due to spatial heterogeneity and aggregation effects.

The representative elementary volume (REV) concept is frequently employed to address scaling challenges. The REV represents the minimum volume at which soil properties can be considered statistically representative of larger systems (34, 52). Determining appropriate REV's remains difficult because soil heterogeneity varies considerably across landscapes and management systems.

Temporal scaling presents additional complexities. Soil properties are dynamic and respond continuously to climatic variability, biological activity, and land management practices. Mechanistic PTFs must therefore account for temporal changes in structure, organic matter content, and hydraulic behavior. Static models may inadequately capture these dynamics, reducing predictive reliability.

Upscaling and downscaling techniques have been proposed to bridge gaps between scales. Upscaling seeks to translate pore-scale information into field-scale predictions, whereas downscaling aims to derive fine-resolution information from broader-scale observations. Hybrid approaches combining mechanistic understanding with statistical and machine-learning methods have shown promise for addressing scaling challenges while maintaining computational efficiency (5).

5.6 Frameworks for Mechanistic Pedotransfer Development

The development of mechanistic pedotransfer functions generally follows a structured framework linking measurable soil attributes to process-based representations and ultimately to target soil properties. Although methodologies vary, most frameworks involve four key stages: data acquisition, process characterization, model formulation, and validation.

The first stage involves collection of soil physical, chemical, biological, and structural data. These datasets may include texture, bulk density, organic carbon content, aggregate stability, pore characteristics, mineralogy, and environmental covariates. Advances in proximal sensing, remote sensing, and imaging technologies have expanded the range of information available for model development (13, 53, 54).

The second stage focuses on identifying the processes controlling the target property. For hydraulic conductivity prediction, relevant processes may include pore connectivity, capillary flow, and preferential transport. For carbon storage prediction, decomposition dynamics, mineral-organic interactions, and aggregation mechanisms may be more important.

The third stage involves formulation of predictive relationships based on physical laws, process-based simulations, or hybrid approaches. Mechanistic equations, pore-network models, hydrological simulations, and reactive transport models are commonly employed. Increasingly, physics-informed machine learning frameworks are being used to integrate mechanistic constraints with data-driven flexibility (12).

The final stage involves model calibration, validation, uncertainty analysis, and performance evaluation. Independent datasets are essential for assessing predictive reliability and transferability. Sensitivity analyses further identify the most influential variables and processes affecting model outputs.

Emerging frameworks increasingly adopt digital soil twins, explainable artificial intelligence, and process-guided machine learning approaches. These innovations represent a convergence of soil physics, computational science, and data analytics that is expected to shape the future of mechanistic pedotransfer modeling.

Table 1
Comparison of Empirical, Statistical, Machine Learning, and Mechanistic Pedotransfer Functions

Attribute	Empirical PTFs	Statistical PTFs	Machine Learning-Based PTFs	Mechanistic PTFs
Fundamental Principle	Direct relationships derived from observed data	Statistical inference from multiple predictor variables	Learning complex patterns from large datasets	Prediction based on physical, chemical, and biological soil processes
Theoretical Foundation	Empirical observations	Statistical theory	Computational learning theory	Soil physics, hydrology, and process-based science
Typical Methods	Simple linear and nonlinear regression	Multiple regression, PCA, Bayesian models, geostatistics	Artificial Neural Networks (ANN), Random Forest (RF), Support Vector Machine (SVM), Gradient Boosting, Deep Learning	Pore-network models, capillary bundle models, Richards-based models, process-based simulations
Input Variables	Texture, bulk density, organic matter	Multiple soil physical and chemical attributes	Soil properties, terrain variables, climate data, remote sensing products	Texture, structure, pore characteristics, hydraulic parameters, environmental drivers
Representation of Soil Processes	Minimal	Limited	Implicit	Explicit
Ability to Capture Nonlinearity	Low	Moderate	High	Moderate to High
Interpretability	High	Moderate to High	Low (often black-box)	High
Data Requirement	Low	Moderate	High	Moderate to High
Computational Demand	Low	Moderate	High	High
Uncertainty Quantification	Limited	Moderate to High	Variable	Moderate to High
Transferability Across Regions	Often Poor	Moderate	Variable	Generally Better

Attribute	Empirical PTFs	Statistical PTFs	Machine Learning-Based PTFs	Mechanistic PTFs
Extrapolation Beyond Calibration Domain	Weak	Moderate	Often Unreliable	Stronger due to physical constraints
Sensitivity to Dataset Bias	High	Moderate	High	Moderate
Scalability	High	High	High	Moderate
Ability to Integrate Multi-Source Data	Limited	Moderate	Excellent	Excellent
Suitability for Dynamic Simulations	Poor	Limited	Moderate	Excellent
Major Strength	Simplicity and ease of application	Improved prediction and uncertainty assessment	High predictive accuracy and flexibility	Physical realism, interpretability, and process representation
Major Limitation	Site-specific applicability and weak extrapolation	Dependence on statistical assumptions	Black-box behavior and overfitting risks	High data requirements and computational complexity
Typical Applications	Estimation of field capacity, wilting point, bulk density	Soil hydraulic characterization, soil surveys	Digital soil mapping, large-scale prediction, precision agriculture	Hydraulic conductivity prediction, water retention modeling, carbon dynamics, Earth system modeling

Table Footnote

PTF = Pedotransfer Function; PCA = Principal Component Analysis; ANN = Artificial Neural Network; RF = Random Forest; SVM = Support Vector Machine. Adapted and synthesized from (3–6, 9, 12, 14).

6. Data Requirements for Mechanistic Pedotransfer Functions

6.1 Conventional Soil Survey Data

Conventional soil survey data constitute the foundational input for most pedotransfer functions and remain indispensable in mechanistic frameworks. These datasets typically include particle-size distribution (sand, silt, and clay contents), bulk density, soil depth, organic carbon content, pH, cation exchange capacity (CEC), electrical conductivity, and horizon characteristics obtained through field surveys and laboratory analyses (3, 6). Because these properties are routinely measured and widely

available, they provide the primary link between observable soil characteristics and functional soil behavior.

In mechanistic pedotransfer functions (MPTFs), conventional soil properties serve not merely as predictor variables but as indicators of underlying physical and pedogenic processes. Soil texture, for example, influences pore-size distribution, hydraulic conductivity, and water retention through its control over particle arrangement and surface area. Similarly, bulk density provides information regarding compaction, porosity, and structural development. Organic carbon content reflects biological activity and aggregate stabilization processes that affect hydraulic and chemical behavior (17).

National and global soil databases such as HYPRES, SoilGrids, ISRIC databases, and national soil survey inventories have significantly expanded opportunities for MPTF development by providing standardized datasets across diverse environmental conditions (13). However, conventional soil data alone often inadequately capture dynamic structural and biological processes, necessitating the integration of additional information sources.

6.2 Soil Morphological and Structural Attributes

Soil morphology and structure provide critical information regarding the organization of the soil matrix and the processes governing water movement, aeration, and root development. Morphological characteristics commonly assessed in field descriptions include soil color, horizon boundaries, consistence, mottling, root abundance, pore distribution, and structural type (55). These attributes offer valuable insights into drainage conditions, redox processes, biological activity, and pedogenic evolution.

Structural attributes are particularly important for mechanistic prediction because they directly influence pore architecture. Aggregate size distribution, aggregate stability, total porosity, macroporosity, pore connectivity, tortuosity, and bulk density collectively determine the movement and retention of water, gases, and solutes (31). Two soils with similar textures may exhibit markedly different hydraulic properties due to variations in structural organization, highlighting the limitations of texture-based prediction alone.

Recent advances in imaging technologies have enabled quantitative characterization of soil structure at unprecedented resolutions. X-ray computed tomography (CT), magnetic resonance imaging (MRI), scanning electron microscopy (SEM), and digital image analysis provide detailed information on pore geometry and connectivity (23). These measurements are increasingly incorporated into mechanistic frameworks to improve prediction of hydraulic conductivity, infiltration, and water retention characteristics.

Because structure reflects both pedogenic history and ongoing biological activity, its incorporation enhances the ability of MPTFs to represent real-world soil behavior under varying environmental conditions.

6.3 Hydraulic Measurements

Hydraulic measurements provide essential information for understanding and predicting soil water dynamics. Since many pedotransfer functions were originally developed to estimate hydraulic properties, hydraulic datasets remain central to mechanistic modeling efforts (9).

Key hydraulic measurements include saturated hydraulic conductivity (K_s), unsaturated hydraulic conductivity, infiltration rate, soil water retention characteristics, field capacity, permanent wilting point, and plant-available water content. These properties describe the capacity of soils to store and transmit water and are fundamental inputs to hydrological, agricultural, and environmental models (24, 37).

The soil water retention curve (SWRC) is among the most important hydraulic datasets used in MPTFs. It characterizes the relationship between soil water content and matric potential, reflecting the influence of pore-size distribution on water storage. Similarly, hydraulic conductivity measurements provide information regarding the ease with which water moves through soil pores under varying moisture conditions.

Mechanistic pedotransfer models often utilize hydraulic measurements during calibration and validation stages. Pore-scale models, capillary bundle theories, and process-based hydrological simulations frequently require hydraulic parameters to establish relationships between pore architecture and water flow dynamics (56, 57). Consequently, high-quality hydraulic measurements are essential for developing physically realistic predictive frameworks.

Emerging technologies such as tension infiltrometers, automated soil moisture sensors, time-domain reflectometry (TDR), and dielectric probes have improved the availability and accuracy of hydraulic data, thereby supporting more sophisticated mechanistic modeling approaches.

6.4 Soil Biological Indicators

The importance of biological processes in soil functioning has increasingly been recognized within mechanistic pedotransfer research. Biological indicators provide information regarding microbial activity, organic matter turnover, nutrient cycling, aggregation processes, and root–soil interactions, all of which influence soil physical and chemical properties (8).

Common biological indicators include microbial biomass carbon, microbial respiration, enzyme activities, earthworm abundance, root density, microbial diversity indices, and soil fauna populations. These indicators reflect the biological drivers responsible for organic matter decomposition, aggregate formation, nutrient mineralization, and carbon stabilization (40).

Microorganisms play a particularly important role in soil structure formation through the production of extracellular polymeric substances and fungal hyphae that bind particles into aggregates. Similarly, earthworms and other soil fauna create biopores that alter hydraulic conductivity and preferential flow pathways (29). Root systems further modify soil architecture through pore creation, exudation, and rhizosphere interactions.

Incorporating biological indicators into mechanistic pedotransfer models enhances the representation of dynamic soil processes that cannot be adequately captured by physical measurements alone. As interest grows in ecosystem services, soil health assessment, and carbon sequestration, biological data are expected to become increasingly important components of next-generation predictive frameworks.

6.5 Environmental Covariates

Environmental covariates provide contextual information that influences soil formation, development, and functioning across landscapes. Within mechanistic pedotransfer modeling, these variables serve as proxies for the environmental drivers controlling pedogenic, hydrological, and biological processes (58).

Climate variables, including precipitation, temperature, evapotranspiration, humidity, and radiation, strongly influence weathering rates, organic matter decomposition, water balance, and vegetation dynamics. Consequently, climatic datasets are frequently incorporated into predictive models to improve representation of soil process variability (25, 26).

Topographic attributes such as elevation, slope gradient, slope position, curvature, aspect, and topographic wetness index influence erosion, deposition, drainage patterns, and microclimatic conditions. Digital elevation models (DEMs) have therefore become important sources of environmental covariates in predictive soil modeling.

Land-use and vegetation information further contribute to mechanistic understanding by reflecting differences in management practices, organic matter inputs, root distributions, and biological activity. Remote sensing products derived from satellite platforms increasingly provide spatially continuous datasets describing vegetation cover, land surface temperature, and ecosystem productivity (13).

By integrating environmental covariates with conventional soil properties, MPTFs can better represent the processes responsible for spatial and temporal variability in soil functions.

6.6 Multi-Scale Data Integration

A defining characteristic of mechanistic pedotransfer functions is the integration of information across multiple spatial, temporal, and organizational scales. Soil processes operate from nanometer-scale mineral interactions to landscape-scale hydrological dynamics, necessitating frameworks capable of linking observations across these scales (51).

At the microscale, imaging technologies provide detailed information regarding pore geometry, particle arrangement, and mineral-organic associations. At the pedon scale, field measurements capture horizon development, hydraulic properties, and biological activity. Landscape-scale datasets derived from remote sensing, digital elevation models, and climatic observations provide broader environmental context.

The integration of multi-scale datasets requires advanced computational approaches capable of handling heterogeneous data types and resolutions. Geographic Information Systems (GIS), digital soil mapping

techniques, data assimilation methods, and machine learning algorithms have become important tools for combining information across scales (14).

Recent developments in digital twins, physics-informed machine learning, and Earth system modeling further emphasize the importance of multi-scale integration. These approaches enable simultaneous representation of local soil processes and broader environmental controls, thereby improving predictive accuracy and transferability (12).

Ultimately, multi-scale data integration represents a critical step toward the development of robust mechanistic pedotransfer functions capable of supporting precision agriculture, hydrological forecasting, climate modeling, and sustainable land management in complex and changing environments.

Table 2
Major Data Requirements and Their Roles in Mechanistic Pedotransfer Functions

Data Category	Examples	Primary Information Provided	Role in Mechanistic PTFs
Conventional Soil Data	Texture, bulk density, SOC, pH, CEC	Basic soil characteristics	Foundation for process representation
Morphological Data	Horizon depth, color, structure, consistence	Pedogenic development	Indicates soil-forming processes
Structural Data	Porosity, aggregate stability, pore connectivity	Soil architecture	Controls hydraulic and transport processes
Hydraulic Data	Ks, SWRC, infiltration rate	Water storage and movement	Calibration and validation of hydraulic predictions
Biological Data	Microbial biomass, respiration, root density	Biological activity	Represents dynamic soil processes
Environmental Covariates	Climate, topography, land use	External controls	Captures environmental variability
Remote Sensing Data	NDVI, land surface temperature	Landscape conditions	Supports spatial prediction
Imaging Data	CT scans, SEM images	Pore geometry and connectivity	Enables pore-scale mechanistic modeling
Multi-Scale Integrated Data	Combined datasets	System-level understanding	Supports comprehensive process-based prediction

Source

Adapted from (5, 6, 9, 14, 58).

7. Process-Based Modeling Approaches in Mechanistic Pedotransfer Functions

7.1 Soil Water Retention Modeling

Soil water retention is one of the most extensively studied properties in pedotransfer research because it regulates plant water availability, infiltration, runoff generation, nutrient transport, and hydrological responses. Mechanistic pedotransfer functions (MPTFs) predict water retention by representing the physical processes governing water storage within soil pore systems rather than relying solely on empirical correlations (9).

The theoretical basis of water retention modeling lies in capillary theory, which relates water retention to pore radius distribution. Water occupies progressively smaller pores as matric potential decreases, resulting in characteristic soil water retention curves (SWRCs). These curves describe the relationship between volumetric water content and matric potential and provide essential parameters for hydrological simulations (21, 33).

Among the most influential process-based retention models are the Brooks–Corey and van Genuchten models. The Brooks–Corey model conceptualizes pore systems using capillary bundle theory and describes water retention through pore-size distributions (59). The van Genuchten model introduced a flexible mathematical formulation capable of accurately representing retention behavior across a wide range of soil conditions and remains widely applied in mechanistic PTF development (60).

Mechanistic approaches increasingly integrate soil structural information, pore-size distributions, and imaging-derived metrics into retention predictions. Recent developments combine pore-scale observations with hydraulic theories to improve model realism and predictive capability under varying environmental conditions.

7.2 Hydraulic Conductivity Prediction

Hydraulic conductivity represents the capacity of soil to transmit water and is among the most important hydraulic properties predicted through pedotransfer functions. Saturated hydraulic conductivity (K_s) and unsaturated hydraulic conductivity govern infiltration, drainage, groundwater recharge, and plant water uptake (48).

Traditional empirical PTFs estimate conductivity directly from texture and bulk density. In contrast, mechanistic approaches derive conductivity from physical representations of pore architecture and fluid flow processes. Darcy's law provides the foundational framework describing saturated flow through porous media, while unsaturated flow is commonly represented through Richards' Eq. (35).

Mechanistic conductivity models often incorporate pore connectivity, tortuosity, and pore-size distribution as primary determinants of flow behavior. The Mualem–van Genuchten framework remains one of the most widely used approaches, linking hydraulic conductivity to water retention characteristics through physically meaningful relationships (60, 61).

Recent advances in computed tomography and pore-network analysis have enabled direct estimation of hydraulic conductivity from three-dimensional pore geometries. These developments have improved understanding of preferential flow pathways, structural heterogeneity, and the effects of biological pores on water movement.

Hydraulic conductivity prediction remains a central application of mechanistic PTFs because it directly connects soil structure to hydrological function and serves as a critical parameter in environmental and agricultural modeling.

7.3 Pore-Scale and Pore-Network Models

Pore-scale and pore-network models represent some of the most physically realistic approaches used in mechanistic pedotransfer research. These models explicitly describe the geometry, connectivity, and topology of pore systems responsible for water and solute movement through soils (62).

Pore-scale models utilize high-resolution imaging techniques such as X-ray computed tomography (CT), scanning electron microscopy (SEM), and synchrotron imaging to reconstruct three-dimensional soil architectures. These reconstructions permit direct simulation of fluid flow and transport processes at the pore level, thereby linking microscopic structure to macroscopic soil properties (23).

Pore-network models simplify complex pore geometries into interconnected networks of pores and throats. Flow and transport are then simulated through these networks using physical laws governing capillary forces, viscous flow, and diffusion. Such models have successfully predicted water retention, hydraulic conductivity, gas diffusion, and solute transport across diverse soil types (63).

One major advantage of pore-network approaches is their ability to represent structural heterogeneity and preferential flow pathways. Unlike continuum-based models that assume uniform flow conditions, pore-network models capture localized processes responsible for rapid transport and non-equilibrium behavior.

Despite their scientific value, pore-scale models remain computationally intensive and often require specialized imaging data. Consequently, efforts continue to develop simplified mechanistic frameworks that retain essential pore-scale information while remaining operationally feasible for large-scale applications.

7.4 Physically Based Soil Hydrological Models

Physically based hydrological models simulate water movement through soils using fundamental conservation laws and process-based equations. These models form an important component of mechanistic pedotransfer frameworks because they provide direct links between measurable soil properties and hydrological behavior (64).

Most hydrological models rely on Richards' equation to simulate unsaturated water flow. This equation combines Darcy's law with mass conservation principles and accounts for changes in hydraulic

conductivity as soil moisture varies. Numerical solutions of Richards' equation form the basis of widely used models such as HYDRUS, SWAP, and SHAW (65).

Hydrological models require hydraulic parameters that are often difficult to measure directly. Mechanistic PTFs therefore play a critical role by providing physically informed estimates of these parameters from readily available soil data. Soil water retention characteristics, hydraulic conductivity functions, and pore-size distributions commonly serve as inputs to hydrological simulations.

Recent developments emphasize coupling hydrological models with climate data, remote sensing observations, and digital soil mapping products. Such integration improves predictions of infiltration, runoff generation, groundwater recharge, evapotranspiration, and drought responses.

The incorporation of physically based hydrological models into pedotransfer frameworks enhances process realism and strengthens the capacity of predictive systems to operate under changing environmental conditions.

7.5 Reactive Transport Models

Reactive transport models integrate physical transport processes with chemical and biological reactions occurring within soil systems. These models are particularly important for predicting nutrient cycling, contaminant behavior, carbon dynamics, and biogeochemical transformations (66).

Transport components describe advection, diffusion, and dispersion processes responsible for the movement of dissolved substances through soils. Simultaneously, reaction modules simulate sorption, precipitation, dissolution, ion exchange, redox reactions, and microbial transformations. The integration of transport and reaction processes provides a comprehensive representation of soil chemical dynamics.

Widely used reactive transport models include HYDRUS-HP1, CrunchFlow, TOUGHREACT, and PHREEQC. These frameworks allow simulation of nutrient availability, contaminant mobility, carbon sequestration processes, and soil acidification under varying environmental conditions (67).

Mechanistic pedotransfer functions increasingly support reactive transport modeling by providing estimates of hydraulic and geochemical parameters that are difficult to measure directly. Examples include adsorption coefficients, pore-water velocities, hydraulic conductivities, and moisture retention characteristics.

As environmental concerns related to pollution, nutrient losses, and climate change intensify, reactive transport models are expected to play an increasingly important role in process-based soil prediction.

7.6 Dynamic Soil Process Models

Dynamic soil process models represent the most comprehensive class of mechanistic frameworks because they simultaneously integrate physical, chemical, and biological processes operating over time. Unlike static models, dynamic approaches explicitly account for temporal changes in soil properties

resulting from management practices, climatic variability, vegetation dynamics, and biological activity (68).

Examples include carbon turnover models such as CENTURY, RothC, and DNDC, which simulate organic matter decomposition, nutrient cycling, and greenhouse gas emissions. These models represent interactions among soil organic matter pools, microbial activity, environmental conditions, and management interventions (69).

Dynamic soil process models are particularly valuable for evaluating long-term changes in soil quality, carbon sequestration potential, and ecosystem resilience. Their integration into mechanistic pedotransfer frameworks enables prediction of soil properties under future climatic scenarios and evolving land-use systems.

Recent advances in Earth system modeling, digital twins, and process-guided machine learning have further expanded the scope of dynamic modeling approaches. These innovations facilitate real-time updating of soil predictions while maintaining consistency with established physical and biogeochemical principles (12).

The future of mechanistic pedotransfer research is likely to involve increasingly sophisticated dynamic frameworks capable of integrating multi-source observations, process understanding, and predictive analytics across multiple spatial and temporal scales.

Table 3
Major Process-Based Models Employed in Mechanistic Pedotransfer Studies

Model	Model Type	Primary Process Represented	Key Input Data	Typical Application
Brooks–Corey	Water retention model	Soil water retention	Texture, pore-size distribution	Estimation of SWRC parameters
van Genuchten	Hydraulic model	Water retention and conductivity	Texture, bulk density, hydraulic data	Hydraulic parameter estimation
Mualem–van Genuchten	Conductivity model	Unsaturated hydraulic conductivity	SWRC parameters	Flow simulation
Richards Equation	Hydrological model	Unsaturated water flow	Hydraulic conductivity, matric potential	Water movement prediction
HYDRUS-1D/2D/3D	Hydrological model	Water, heat, and solute transport	Hydraulic properties, climate data	Vadose zone simulation
SWAP	Soil-water-atmosphere model	Soil water balance	Climate, crop, soil data	Agricultural water management
SHAW	Coupled heat-water model	Heat and water transport	Meteorological and soil data	Soil–plant–atmosphere interactions
Pore-Network Models	Pore-scale model	Flow through pore systems	CT-derived pore geometry	Hydraulic conductivity prediction
Lattice Boltzmann Models	Pore-scale simulation	Fluid flow in porous media	3D pore images	Microscale transport analysis
PHREEQC	Reactive transport model	Geochemical reactions	Chemical composition data	Nutrient and contaminant modeling
CrunchFlow	Reactive transport model	Multicomponent transport and reactions	Hydraulic and geochemical parameters	Soil and groundwater studies
TOUGHREACT	Reactive transport model	Coupled transport-reaction processes	Physical and chemical data	Environmental simulations
CENTURY	Dynamic ecosystem model	Carbon and nutrient cycling	Climate, soil, vegetation data	Long-term soil change prediction

Model	Model Type	Primary Process Represented	Key Input Data	Typical Application
RothC	Carbon dynamics model	Soil organic carbon turnover	SOC, climate data	Carbon sequestration studies
DNDC	Biogeochemical model	Carbon and nitrogen cycling	Climate, management, soil data	Greenhouse gas assessments

Source

Adapted and synthesized from (12, 59, 60, 65, 66).

8. Integration of Data-Driven and Mechanistic Modeling

8.1 Hybrid Modeling Concepts

The increasing availability of large soil datasets, advances in computational power, and growing demand for accurate soil predictions have stimulated the development of hybrid modeling approaches that combine mechanistic understanding with data-driven techniques. Hybrid models seek to leverage the strengths of both paradigms: the interpretability and physical realism of process-based models and the predictive flexibility of machine learning algorithms (11, 53, 54).

Traditional pedotransfer functions often occupy opposite ends of the modeling spectrum. Purely empirical and machine learning models generally provide strong predictive performance but may lack physical interpretability and struggle when extrapolated beyond calibration domains. Conversely, mechanistic models offer causal understanding but frequently require extensive parameterization and computational resources. Hybrid approaches attempt to bridge this divide by embedding physical principles into data-driven frameworks or using machine learning to enhance mechanistic simulations (12).

Several hybrid strategies have emerged in soil science. These include residual correction approaches, where machine learning models correct errors generated by mechanistic simulations; surrogate modeling, where computationally expensive process-based models are approximated using data-driven algorithms; and coupled frameworks that simultaneously integrate physical equations and machine learning architectures (70).

In mechanistic pedotransfer research, hybrid models have demonstrated considerable potential for predicting hydraulic conductivity, soil water retention, organic carbon dynamics, and nutrient transport. By combining observational data with process understanding, these models improve prediction accuracy while maintaining scientific plausibility.

8.2 Machine Learning-Augmented Mechanistic Pedotransfer Functions

Machine learning-augmented mechanistic pedotransfer functions represent an important advancement in predictive soil science. Rather than replacing mechanistic models, machine learning algorithms are employed to complement and enhance process-based predictions. This approach acknowledges that while physical models capture known processes, they may not fully represent the complexity and heterogeneity of natural soil systems.

One common application involves estimating difficult-to-measure model parameters. For example, machine learning algorithms can predict hydraulic parameters required by physically based hydrological models using readily available soil survey data. Artificial neural networks (ANNs), random forests (RFs), gradient boosting machines (GBMs), and support vector machines (SVMs) have been successfully applied for this purpose (14).

Another strategy involves residual learning, whereby machine learning models predict discrepancies between mechanistic model outputs and observed measurements. These residual corrections can substantially improve model accuracy without sacrificing physical consistency. Such approaches have proven effective in hydrological forecasting, carbon dynamics modeling, and digital soil mapping (70).

Machine learning can also assist in identifying nonlinear relationships and interactions among soil variables that may be overlooked by traditional mechanistic models. By integrating these insights into process-based frameworks, researchers can develop more robust pedotransfer functions capable of representing complex soil behavior.

Despite their advantages, machine learning-augmented models require careful validation to avoid overfitting and ensure that predictions remain consistent with established physical principles. Consequently, the integration of machine learning and mechanistic understanding must be guided by both statistical rigor and domain expertise.

8.3 Physics-Informed Machine Learning

Physics-informed machine learning (PIML) represents one of the most promising developments in the integration of data science and process-based modeling. Unlike conventional machine learning models that learn solely from data, PIML incorporates physical laws, conservation principles, and mechanistic constraints directly into the learning process (71).

In soil science, physical constraints may include Darcy's law, Richards' equation, mass conservation, energy balance equations, and biogeochemical reaction kinetics. By embedding these principles into model architectures or loss functions, PIML ensures that predictions remain physically realistic even when observational data are limited.

Physics-informed neural networks (PINNs) have emerged as a particularly influential class of PIML methods. These networks integrate differential equations into neural network training, enabling simultaneous learning from observations and governing physical laws. PINNs have been applied successfully to groundwater flow simulation, unsaturated water transport, and soil moisture prediction (72).

For mechanistic pedotransfer functions, physics-informed learning offers several advantages. First, it reduces dependence on large training datasets by leveraging prior scientific knowledge. Second, it improves extrapolation performance under novel environmental conditions. Third, it enhances interpretability by ensuring that model behavior aligns with established process understanding.

The adoption of PIML is expected to accelerate as soil scientists seek predictive frameworks capable of integrating heterogeneous datasets while maintaining physical credibility and scientific transparency.

8.4 Digital Twins and Soil System Modeling

Digital twin technology represents a transformative development in environmental and agricultural sciences. A digital twin is a dynamic virtual representation of a physical system that continuously updates using real-time observations, simulations, and predictive analytics (73). In soil science, digital twins offer opportunities to create comprehensive representations of soil systems that integrate physical, chemical, biological, and hydrological processes.

Soil digital twins combine field measurements, remote sensing observations, sensor networks, process-based models, and machine learning algorithms to simulate soil behavior across multiple spatial and temporal scales. These systems can continuously assimilate new information, enabling real-time monitoring and prediction of soil conditions.

Mechanistic pedotransfer functions serve as important components within soil digital twins because they provide links between observable soil attributes and functional properties. Hydraulic conductivity, soil moisture retention, carbon stocks, nutrient availability, and biological activity can all be estimated through integrated pedotransfer frameworks and updated as new data become available.

Digital twins also facilitate scenario analysis by allowing researchers and land managers to evaluate the impacts of climate change, land-use modifications, irrigation practices, and conservation strategies before implementation. Such capabilities support precision agriculture, sustainable land management, and environmental decision-making.

As sensor technologies, cloud computing, and artificial intelligence continue to advance, digital twin systems are expected to become increasingly important platforms for integrating mechanistic and data-driven soil predictions.

8.5 Explainable Artificial Intelligence in Mechanistic Prediction

One of the primary criticisms of machine learning approaches in soil science is their "black-box" nature, whereby accurate predictions are generated without clear explanations of the underlying processes. Explainable Artificial Intelligence (XAI) seeks to address this limitation by providing interpretable insights into model behavior, variable importance, and prediction pathways (74, 75).

In mechanistic pedotransfer modeling, explainability is particularly important because predictive performance alone is insufficient. Soil scientists require models that not only generate accurate estimates but also enhance understanding of soil processes. XAI techniques help bridge this gap by revealing how predictors contribute to model outputs and whether predictions align with established scientific knowledge.

Common XAI methods include feature importance analysis, partial dependence plots, Shapley Additive Explanations (SHAP), Local Interpretable Model-Agnostic Explanations (LIME), and sensitivity analyses. These approaches allow researchers to identify dominant controls on soil properties and evaluate whether model predictions are physically plausible (76).

The integration of XAI with mechanistic pedotransfer functions offers several advantages. First, it enhances model transparency and user confidence. Second, it facilitates scientific discovery by identifying previously unrecognized relationships among soil variables. Third, it supports model validation by ensuring consistency between machine learning outputs and established process understanding.

Future soil prediction frameworks are likely to combine mechanistic models, machine learning algorithms, and explainability tools within unified systems that provide both predictive accuracy and scientific insight. Such developments represent a critical step toward trustworthy artificial intelligence in soil and environmental sciences.

Table 4

Emerging Approaches for Integrating Data-Driven and Mechanistic Modeling in Pedotransfer Functions

Approach	Core Principle	Typical Techniques	Advantages	Limitations	Applications in Soil Science
Hybrid Modeling	Combines process-based and data-driven methods	Coupled models, residual correction	Improved accuracy and realism	Increased complexity	Hydraulic property prediction
Machine Learning-Augmented Mechanistic Models	ML enhances mechanistic simulations	ANN, RF, SVM, GBM	Captures nonlinear interactions	Risk of overfitting	Hydraulic conductivity and carbon prediction
Physics-Informed Machine Learning	Physical laws embedded into learning algorithms	PINNs, constrained optimization	Improved extrapolation and consistency	Computationally intensive	Water flow and transport modeling
Digital Twins	Real-time virtual representation of soil systems	Sensors, AI, process models	Continuous updating and forecasting	High data requirements	Precision agriculture and soil monitoring
Explainable AI	Interpretation of model predictions	SHAP, LIME, PDPs	Transparency and trustworthiness	Additional computational cost	Mechanistic model interpretation
Process-Guided Machine Learning	Process knowledge guides ML architecture	Hybrid neural networks	Balances realism and flexibility	Model design complexity	Earth system and soil process modeling

Source

Adapted from (11, 12, 70, 71, 74).

9. Applications of Mechanistic Pedotransfer Functions

9.1 Prediction of Soil Hydraulic Properties

The prediction of soil hydraulic properties remains the most prominent application of mechanistic pedotransfer functions (MPTFs). Hydraulic properties such as soil water retention, field capacity, permanent wilting point, saturated hydraulic conductivity, and unsaturated hydraulic conductivity are essential inputs for hydrological, agricultural, and environmental models (6). Direct measurement of these properties is often labor-intensive, expensive, and time-consuming, creating a need for reliable predictive approaches.

Mechanistic PTFs improve hydraulic property estimation by explicitly representing pore-size distribution, capillary behavior, pore connectivity, and water flow processes. Models based on capillary bundle theory, pore-network representations, and hydraulic conductivity functions provide physically meaningful estimates that are generally more transferable than purely empirical models (9). The integration of structural and hydraulic information has significantly enhanced the prediction of water retention curves and conductivity functions across diverse soil types.

Recent developments incorporating X-ray computed tomography, image analysis, and process-based hydrological simulations have further strengthened hydraulic predictions. As a result, MPTFs have become important tools for supporting irrigation design, drainage management, watershed modeling, and groundwater recharge assessments.

9.2 Soil Water Management and Irrigation Planning

Efficient water management is fundamental to sustainable agriculture, particularly in regions facing increasing water scarcity and climatic variability. Mechanistic pedotransfer functions contribute to irrigation planning by providing estimates of soil hydraulic properties that regulate water storage, infiltration, redistribution, and plant water availability (24, 33).

By predicting field capacity, permanent wilting point, and plant-available water content, MPTFs enable more accurate determination of irrigation requirements. These estimates assist in optimizing irrigation scheduling, reducing water losses, and improving crop water-use efficiency. Unlike empirical models, mechanistic approaches provide insights into the processes governing water movement, allowing more reliable predictions under changing environmental conditions.

Mechanistic predictions are also integrated into soil–water–atmosphere models used to evaluate irrigation strategies under different climate scenarios. Such applications are particularly relevant in semi-arid and arid regions where efficient allocation of water resources is essential for agricultural sustainability (77).

The increasing adoption of sensor technologies and precision irrigation systems further enhances the value of MPTFs by enabling real-time estimation of soil water dynamics and irrigation requirements.

9.3 Land Degradation Assessment

Land degradation, including erosion, compaction, salinization, nutrient depletion, and desertification, represents a major threat to global food security and ecosystem sustainability. Mechanistic pedotransfer functions contribute to degradation assessment by predicting soil properties that serve as indicators of soil health and functional capacity (78).

Changes in soil structure, porosity, infiltration capacity, hydraulic conductivity, and organic matter content are often early indicators of degradation processes. Mechanistic models provide process-based estimates of these properties, allowing researchers to identify degradation trends before substantial productivity losses occur.

For example, reductions in pore connectivity and aggregate stability can be linked to compaction and erosion processes, while declining water retention capacity may indicate structural deterioration. Mechanistic frameworks facilitate the assessment of such changes by relating measurable soil attributes to underlying degradation mechanisms.

In addition, MPTFs support large-scale monitoring programs by integrating field observations, remote sensing data, and environmental covariates. This capability improves the identification of vulnerable areas and informs land restoration and conservation strategies.

9.4 Soil Carbon Modeling

Soil organic carbon (SOC) plays a critical role in soil fertility, nutrient cycling, aggregate stability, and climate regulation. Mechanistic pedotransfer functions increasingly support soil carbon modeling by linking measurable soil properties to carbon storage and turnover processes (40).

Carbon dynamics are governed by complex interactions among organic matter inputs, microbial decomposition, mineral stabilization, environmental conditions, and land management practices. Mechanistic approaches incorporate these processes into predictive frameworks, improving estimation of carbon stocks and sequestration potential.

The integration of MPTFs into carbon models such as RothC, CENTURY, and DNDC enhances the representation of soil physical controls on decomposition and stabilization processes. For example, predictions of pore-size distribution and aggregate structure can improve understanding of physical protection mechanisms that regulate carbon persistence.

As international initiatives increasingly emphasize carbon sequestration for climate mitigation, accurate prediction of SOC dynamics has become a major application area for mechanistic pedotransfer research.

9.5 Climate Change Impact Studies

Climate change is expected to alter temperature regimes, precipitation patterns, evapotranspiration rates, and vegetation dynamics, thereby affecting soil processes and ecosystem functioning. Mechanistic pedotransfer functions play a critical role in assessing these impacts because they provide physically based predictions capable of operating under novel environmental conditions (5).

Traditional empirical PTFs often perform poorly when applied beyond the conditions under which they were developed. Mechanistic models, however, incorporate physical principles governing water flow, carbon cycling, and nutrient dynamics, enabling more reliable extrapolation to future climates.

Applications include simulation of drought impacts on soil hydraulic properties, assessment of carbon sequestration potential under changing climatic conditions, evaluation of erosion risks associated with extreme rainfall events, and prediction of shifts in soil moisture regimes. These analyses support climate adaptation strategies and sustainable land management planning.

The integration of mechanistic pedotransfer functions with Earth system models further enhances the capacity to evaluate feedbacks between soils, vegetation, hydrology, and climate at regional and global scales.

9.6 Precision Agriculture

Precision agriculture seeks to optimize resource use through site-specific management of soils, water, nutrients, and crops. Mechanistic pedotransfer functions contribute significantly to this objective by providing spatially explicit predictions of soil functional properties that are difficult to measure directly (79).

High-resolution estimates of hydraulic conductivity, available water capacity, nutrient retention, and soil organic carbon enable the development of management zones tailored to local soil conditions. Such information supports variable-rate irrigation, fertilizer application, and tillage practices designed to maximize productivity while minimizing environmental impacts.

The integration of MPTFs with proximal sensing technologies, remote sensing platforms, and geographic information systems (GIS) has enhanced the capacity for real-time soil assessment. Machine learning and digital agriculture platforms increasingly utilize mechanistic predictions to support decision-making and optimize crop management.

As precision agriculture continues to evolve, mechanistic pedotransfer functions are expected to play a central role in translating soil information into actionable management recommendations.

9.7 Digital Soil Mapping and Soil Information Systems

Digital Soil Mapping (DSM) has emerged as a major framework for generating spatially continuous soil information using statistical, machine learning, and process-based methods. Mechanistic pedotransfer functions enhance DSM by providing physically meaningful predictions of soil functional properties from available environmental and soil datasets (58).

Within digital soil information systems, MPTFs are used to estimate hydraulic parameters, carbon stocks, nutrient availability, and other functional attributes that are rarely measured directly at large scales. These estimates contribute to national and global soil databases, including SoilGrids and GlobalSoilMap initiatives.

The integration of remote sensing data, digital elevation models, climate information, and field observations enables the generation of detailed soil property maps across diverse landscapes. Mechanistic approaches improve the reliability and interpretability of these predictions by ensuring consistency with established soil processes.

Future soil information systems are expected to incorporate digital twins, real-time monitoring networks, and physics-informed machine learning models, creating dynamic platforms capable of continuously updating soil property predictions.

Table 5
Selected Applications of Mechanistic Pedotransfer Functions Across Soil Science Disciplines

Application Area	Predicted Soil Properties	Mechanistic Basis	Typical Models/Tools	Major Benefits
Soil Hydrology	Water retention, hydraulic conductivity, infiltration	Capillary flow and pore connectivity	van Genuchten, HYDRUS, Richards-based models	Improved water flow prediction
Irrigation Management	Field capacity, wilting point, available water	Soil water balance processes	SWAP, HYDRUS	Efficient irrigation scheduling
Land Degradation Assessment	Porosity, aggregate stability, infiltration	Structural degradation processes	Erosion and hydrological models	Early degradation detection
Soil Carbon Modeling	SOC stocks, carbon turnover	Decomposition and stabilization mechanisms	CENTURY, RothC, DNDC	Carbon sequestration assessment
Climate Change Studies	Soil moisture, carbon dynamics, nutrient cycling	Coupled soil–climate interactions	Earth System Models	Climate impact evaluation
Precision Agriculture	Water availability, nutrient retention	Site-specific soil functioning	GIS-based decision systems	Resource optimization
Digital Soil Mapping	Hydraulic and chemical properties	Soil-landscape relationships	SoilGrids, DSM platforms	Large-scale prediction
Nutrient Management	Nutrient transport and retention	Reactive transport processes	PHREEQC, HYDRUS-HP1	Improved fertilizer management
Groundwater Protection	Solute transport and leaching	Flow and transport mechanisms	Reactive transport models	Pollution risk assessment
Ecosystem Modeling	Water, carbon, and nutrient fluxes	Integrated process representation	Dynamic ecosystem models	Ecosystem sustainability analysis

Source

Adapted from (5, 6, 9, 12, 58).

10. Uncertainty Analysis and Model Evaluation

10.1 Sources of Uncertainty

Uncertainty is an inherent component of mechanistic pedotransfer modeling because soil systems are highly heterogeneous, dynamic, and influenced by numerous interacting processes. Understanding the sources of uncertainty is essential for improving model reliability and supporting informed decision-making. Uncertainties in mechanistic pedotransfer functions generally arise from data limitations, model structural assumptions, parameter estimation procedures, scale mismatches, and measurement errors (80).

Input data uncertainty is often the largest contributor to prediction errors. Variability in soil texture measurements, bulk density estimates, hydraulic properties, and environmental covariates can significantly influence model outputs. Spatial heterogeneity further complicates data collection because measurements obtained at one location may not adequately represent surrounding areas.

Model structural uncertainty arises when the conceptual representation of soil processes differs from actual system behavior. Even physically based models involve simplifying assumptions regarding pore geometry, flow pathways, biological activity, and chemical reactions. These simplifications can introduce systematic biases into predictions.

Parameter uncertainty results from difficulties in estimating model coefficients and process parameters. Many mechanistic models require hydraulic, geochemical, or biological parameters that are difficult to measure directly, leading to reliance on estimated values that may contain substantial uncertainty.

10.2 Parameter Sensitivity Analysis

Sensitivity analysis evaluates the influence of input variables and model parameters on prediction outcomes. It is an essential component of mechanistic pedotransfer development because it identifies the factors that exert the greatest control on model performance and helps prioritize data collection efforts (81).

Local sensitivity analysis examines the effect of small variations in individual parameters while holding other variables constant. Although computationally efficient, this approach may not adequately capture interactions among parameters. Global sensitivity analysis addresses this limitation by simultaneously evaluating parameter variations across their full ranges.

Common sensitivity analysis techniques include the Morris method, variance-based Sobol analysis, Monte Carlo simulations, and regression-based approaches. These methods provide quantitative measures of parameter importance and uncertainty contributions.

In mechanistic hydraulic models, parameters describing pore-size distribution, bulk density, porosity, and hydraulic conductivity often exhibit high sensitivity. Similarly, carbon dynamics models may be highly sensitive to decomposition rates, microbial activity coefficients, and environmental controls. Understanding these sensitivities improves model calibration and reduces predictive uncertainty.

10.3 Model Validation Techniques

Model validation is the process of evaluating predictive performance against independent observations. It provides evidence regarding the reliability, robustness, and applicability of mechanistic pedotransfer functions under varying environmental conditions (82).

Validation approaches typically involve dividing available datasets into calibration and testing subsets. Independent validation datasets are preferred because they provide unbiased assessments of model performance. Cross-validation techniques are also commonly used when data availability is limited.

Performance metrics frequently employed include the coefficient of determination (R^2), root mean square error (RMSE), mean absolute error (MAE), Nash–Sutcliffe efficiency (NSE), bias statistics, and Kling–Gupta efficiency (KGE). These indicators quantify agreement between observed and predicted values and facilitate comparison among alternative models.

Beyond statistical evaluation, process validation is increasingly emphasized in mechanistic modeling. This approach assesses whether internal model behavior remains consistent with established scientific understanding. A model that generates accurate predictions but unrealistic process representations may have limited applicability under changing conditions.

10.4 Error Propagation in Mechanistic Pedotransfer Functions

Error propagation refers to the transmission and amplification of uncertainties throughout a modeling framework. Because mechanistic pedotransfer functions often serve as inputs to larger hydrological, agricultural, or climate models, understanding error propagation is essential for reliable prediction (83).

Errors introduced during data collection, parameter estimation, or model formulation can propagate through successive modeling stages and influence final outputs. For example, uncertainty in hydraulic conductivity estimates may affect predictions of infiltration, soil moisture dynamics, nutrient transport, and crop water availability.

Monte Carlo simulations are widely used to evaluate error propagation by repeatedly sampling uncertain parameters and generating distributions of model outputs. Bayesian approaches further allow probabilistic characterization of uncertainties and prediction intervals.

Quantifying error propagation helps identify critical sources of uncertainty and supports the development of more robust predictive frameworks. It also provides users with realistic estimates of confidence in model predictions.

10.5 Benchmarking Against Conventional Pedotransfer Functions

Benchmarking involves comparing mechanistic pedotransfer functions with empirical, statistical, and machine learning alternatives to evaluate relative strengths and weaknesses. Such comparisons are

essential for demonstrating the added value of mechanistic approaches and guiding future model development (5).

Empirical and statistical PTFs often perform well within calibration datasets but may exhibit limited transferability across regions and environmental conditions. Machine learning approaches frequently achieve higher predictive accuracy but may lack interpretability and physical consistency. Mechanistic PTFs seek to overcome these limitations through explicit representation of soil processes.

Benchmarking studies typically compare models using common datasets and standardized performance metrics. In addition to predictive accuracy, evaluations often consider interpretability, computational requirements, extrapolation capacity, uncertainty characterization, and data demands.

Recent studies indicate that hybrid mechanistic–machine learning approaches frequently outperform both purely mechanistic and purely data-driven models, suggesting that future pedotransfer development will likely emphasize integration rather than competition between modeling paradigms. Ultimately, benchmarking provides a critical framework for assessing progress toward more reliable, transparent, and transferable soil property prediction systems.

11. Emerging Innovations and Technological Advances

Recent advances in soil sensing, computational science, and artificial intelligence are rapidly transforming mechanistic pedotransfer modeling. High-resolution imaging techniques such as X-ray computed tomography (CT) now enable direct visualization of soil pore architecture, providing unprecedented insights into pore connectivity and hydraulic behavior (23). Proximal soil sensing technologies, including visible–near infrared spectroscopy, electromagnetic induction, and gamma radiometry, facilitate rapid acquisition of soil information at multiple scales. The integration of remote sensing platforms further enhances spatial prediction by providing continuous environmental observations. Simultaneously, high-performance computing and cloud-based infrastructures support large-scale simulations that were previously computationally prohibitive. Emerging concepts such as digital soil twins enable real-time representation of soil systems through continuous assimilation of observational data. Furthermore, physics-informed neural networks and generative artificial intelligence offer new opportunities to combine process understanding with predictive analytics, potentially revolutionizing next-generation soil property prediction and decision-support systems (12, 71).

12. Challenges and Limitations

Despite significant progress, several challenges continue to constrain the development and application of mechanistic pedotransfer functions. Data scarcity remains a major limitation, particularly for hydraulic, structural, and biological parameters required by process-based models. Soil systems are inherently heterogeneous, involving complex interactions among physical, chemical, and biological processes that are difficult to represent comprehensively. Spatial and temporal scaling issues further complicate model development because processes observed at pore scales may not directly translate to field or landscape

scales. Computational requirements associated with pore-scale simulations, reactive transport models, and dynamic ecosystem models can also limit operational applications. Additionally, model transferability across contrasting climatic and edaphic conditions remains uncertain. Biological processes, including microbial dynamics and rhizosphere interactions, are often inadequately represented despite their importance to soil functioning. Finally, the lack of standardized datasets, protocols, and reproducible workflows continues to hinder model comparison, validation, and broader adoption within the soil science community (5).

13. Future Perspectives

Future research is expected to move toward universal mechanistic pedotransfer frameworks capable of integrating physical laws, biological processes, and environmental controls across diverse soil systems. The convergence of mechanistic modeling with digital soil mapping, remote sensing, and multi-source data fusion will likely improve predictive accuracy and spatial applicability. Advances in sensor technologies may facilitate real-time soil monitoring networks that continuously update model predictions and support adaptive land management. Increasing adoption of physics-informed machine learning and explainable artificial intelligence will further strengthen the synergy between process understanding and predictive performance. Digital soil twins are anticipated to emerge as powerful platforms for integrating observations, simulations, and management scenarios within unified decision-support systems. Ultimately, the combination of mechanistic knowledge and advanced computational tools offers substantial potential to improve soil resource assessment, enhance climate resilience, support sustainable land management, and contribute to global food security in an era of environmental change.

14. Conclusions

Mechanistic pedotransfer functions represent a significant advancement in predictive soil science by bridging the gap between empirical prediction and process-based understanding. Unlike traditional pedotransfer approaches, mechanistic frameworks explicitly incorporate the physical, chemical, and biological processes that govern soil behavior, thereby enhancing interpretability, transferability, and predictive robustness. The integration of pore-scale characterization, hydrological modeling, reactive transport simulations, and dynamic ecosystem processes has expanded the scope of soil property prediction beyond conventional statistical relationships. Recent developments in machine learning, physics-informed modeling, digital twins, and advanced sensing technologies further demonstrate the potential of hybrid mechanistic frameworks to address emerging scientific and management challenges. Although issues related to data availability, scaling, computational demands, and model uncertainty remain significant, continued technological innovation and interdisciplinary collaboration are expected to accelerate progress. Mechanistic pedotransfer functions therefore constitute a critical foundation for future soil information systems, precision agriculture, climate adaptation strategies, and sustainable management of soil resources worldwide.

Declarations

ACKNOWLEDGEMENT

I sincerely appreciate Prof. Segun Aruleba, for his valuable guidance, scholarly suggestions, and unwavering support throughout the preparation of this paper. His mentorship played a key role in shaping the quality and focus of this work.

COMPETING INTEREST

The author declares no competing interests.

CONSENT TO PUBLISH DECLARATION

Not applicable

ETHICS AND CONSENT TO PARTICIPATE DECLARATIONS

Not applicable

AUTHOR'S CONTRIBUTION DECLARATION

The author solely conceptualized the topic, conducted the literature search, critically analyzed relevant publications, organized the manuscript structure, and wrote the entire paper. All aspects of this paper were independently completed by the author.

FUNDING DECLARATION

There is no funding received for this work.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

1. Minasny B, Finke P, Stockmann U, Vanwalleghem T, McBratney AB. Resolving the integral connection between pedogenesis and landscape evolution. *Earth Sci Rev.* 2015;150:102–20.
2. Vereecken H, Amelung W, Bauke SL, Bogaen H, Brüggemann N, Montzka C, et al. Soil hydrology in the Earth system. *Nat Reviews Earth Environ.* 2022;3(9):573–87.
3. Bouma J. Using soil survey data for quantitative land evaluation. *Adv Soil Sci.* 1989;9(1989):177–213.
4. McBratney AB, Minasny B, Cattle SR, Vervoort RW. From pedotransfer functions to soil inference systems. *Geoderma.* 2002;109(1–2):41–73.

5. Van Looy K, Bouma J, Herbst M, Koestel J, Minasny B, Mishra U, et al. Pedotransfer functions in Earth system science: Challenges and perspectives. *Rev Geophys*. 2017;55(4):1199–256.
6. Wösten J, Pachepsky YA, Rawls W. Pedotransfer functions: bridging the gap between available basic soil data and missing soil hydraulic characteristics. *J Hydrol*. 2001;251(3–4):123–50.
7. Nguyen PM, Haghverdi A, De Pue J, Botula Y-D, Le KV, Waegeman W, et al. Comparison of statistical regression and data-mining techniques in estimating soil water retention of tropical delta soils. *Biosyst Eng*. 2017;153:12–27.
8. Young IM, Crawford JW. Interactions and self-organization in the soil-microbe complex. *Science*. 2004;304(5677):1634–7.
9. Vereecken H, Weynants M, Javaux M, Pachepsky Y, Schaap M, Genuchten MT. Using pedotransfer functions to estimate the van Genuchten–Mualem soil hydraulic properties: A review. *Vadose Zone J*. 2010;9(4):795–820.
10. Guadagnini A, Neuman SP, Schaap MG, Riva M, editors. Statistical and Scaling Analyses of Neural Network Soil Property Inputs/Outputs at an Arizona Field Site. Special Session on Applications of Modeling and Simulation to Climatic Change and Environmental Sciences. SCITE; 2013.
11. Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, et al. Deep learning and process understanding for data-driven Earth system science. *Nature*. 2019;566(7743):195–204.
12. Shen X, Ma J, Li Y, Li Y, Xia X. The effects of multiple global change factors on soil nutrients across China: a meta-analysis. *Int J Environ Res Public Health*. 2022;19(22):15230.
13. Arrouays D, Lagacherie P, Hartemink AE. Digital soil mapping across the globe. Elsevier; 2017. pp. 1–4.
14. Padarian J, Minasny B, McBratney AB. Using deep learning for digital soil mapping. *Soil*. 2019;5(1):79–89.
15. Fatichi S, Or D, Walko R, Vereecken H, Young MH, Ghezzehei TA, et al. Soil structure is an important omission in Earth System Models. *Nat Commun*. 2020;11(1):522.
16. Rawls W, Gish T, Brakensiek D. Estimating soil water retention from soil physical properties and characteristics. *Advances in Soil Science: Volume 16*: Springer; 1991. pp. 213–34.
17. Minasny B, McBratney A. Limited effect of organic matter on soil available water capacity. *Eur J Soil Sci*. 2018;69(1):39–47.
18. Gupta S, Larson W. A model for predicting packing density of soils using particle-size distribution. *Soil Sci Soc Am J*. 1979;43(4):758–64.
19. Rawls W, Brakensiek D, editors. An infiltration based runoff model for use with a design store. *Hydraulics and Hydrology in the Small Computer Age*. ASCE; 1985.
20. Saxton K, Rawls WJ, Romberger JS, Papendick R. Estimating generalized soil-water characteristics from texture. *Soil Sci Soc Am J*. 1986;50(4):1031–6.
21. Wösten J, Lilly A, Nemes A, Le Bas C. Development and use of a database of hydraulic properties of European soils. *Geoderma*. 1999;90(3–4):169–85.

22. Nemes A, Rawls W, Pachepsky YA, van Genuchten MT. Sensitivity analysis of the nonparametric nearest neighbor technique to estimate soil water retention. *Vadose Zone J.* 2006;5(4):1222–35.
23. Vogel H-J, Weller U, Schlüter S. Quantification of soil structure based on Minkowski functions. *Comput Geosci.* 2010;36(10):1236–45.
24. Giudici M. Modeling water flow in variably saturated porous soils and alluvial sediments. *Sustainability.* 2023;15(22):15723.
25. Jenny H. *Factors of soil formation.* 281 pp. New York. 1941;801.
26. Jenny H. *Factors of soil formation: a system of quantitative pedology.* Courier Corporation; 1994.
27. Brady NC, Weil RR, Weil RR. *The nature and properties of soils.* Prentice Hall Upper Saddle River, NJ; 2008.
28. Simonson RW. Outline of a generalized theory of soil genesis. *Soil Sci Soc Am J.* 1959;23(2):152–6.
29. Six J, Bossuyt H, Degryze S, Denef K. A history of research on the link between (micro) aggregates, soil biota, and soil organic matter dynamics. *Soil Tillage Res.* 2004;79(1):7–31.
30. Sanchez PA. *Properties and Management of Soils in the Tropics.* Cambridge University Press; 2019.
31. Bronick CJ, Lal R. Soil structure and management: a review. *Geoderma.* 2005;124(1–2):3–22.
32. Luxmoore R, Jardine P, Wilson G, Jones J, Zelazny L. Physical and chemical controls of preferred path flow through a forested hillslope. *Geoderma.* 1990;46(1–3):139–54.
33. Tuller M, Or D, Hillel D. Retention of water in soil and the soil water characteristic curve. *Encyclopedia Soils Environ.* 2004;4:278–89.
34. Assouline S, Or D. Conceptual and parametric representation of soil hydraulic properties: A review. *Vadose Zone Journal.* 2013;12(4):vzj2013. 07.0121.
35. Richards LA. Capillary conduction of liquids through porous mediums. *physics.* 1931;1(5):318–33.
36. Ghas SR, Jarvis GT. Mantle flow reversals in cylindrical Earth models. *Phys Earth Planet Inter.* 2007;165(3–4):194–207.
37. Šimůnek J, Köhne JM, Kodešová R, Šejna M. Simulating non equilibrium movement of water, solutes, and particles using HYDRUS: A review of recent applications. *Soil Water Res.* 2008;3(1):S42–51.
38. Cuadros J. Clay minerals interaction with microorganisms: a review. *Clay Miner.* 2017;52(2):235–61.
39. Sparks RSJ. Forecasting volcanic eruptions. *Earth Planet Sci Lett.* 2003;210(1–2):1–15.
40. Lehmann J, Kleber M. The contentious nature of soil organic matter. *Nature.* 2015;528(7580):60–8.
41. Cotrufo MF, Ranalli MG, Haddix ML, Six J, Lugato E. Soil carbon storage informed by particulate and mineral-associated organic matter. *Nat Geosci.* 2019;12(12):989–94.
42. Wu S, Konhauser KO, Chen B, Huang L. Reactive Mineral Sink drives soil organic matter dynamics and stabilization. *Npj Mater Sustain.* 2023;1(1):3.
43. Lal R. Sustainable intensification of China's agroecosystems by conservation agriculture. *Int Soil Water Conserv Res.* 2018;6(1):1–12.
44. Lynch JP. Root phenotypes for improved nutrient capture: an underexploited opportunity for global agriculture. *New Phytol.* 2019;223(2):548–64.

45. Jones DL, Nguyen C, Finlay RD. Carbon flow in the rhizosphere: carbon trading at the soil–root interface. Springer; 2009.
46. Smet S, Beckers E, Plougonven E, Léonard A, Degré A. Can the pore scale geometry explain soil sample scale hydrodynamic properties? *Front Environ Sci*. 2018;6:20.
47. Peters A, Hohenbrink TL, Iden SC, Van Genuchten MT, Durner W. Prediction of the absolute hydraulic conductivity function from soil water retention data. *Hydrol Earth Syst Sci*. 2023;27(7):1565–82.
48. Jury WA, Horton R. Soil physics: Wiley; 2004.
49. Strawn DG. Sorption mechanisms of chemicals in soils. *Soil Syst*. 2021;5(1):13.
50. Jarvis N, Larsbo M, Roulier S, Lindahl A, Persson L. The role of soil properties in regulating non-equilibrium macropore flow and solute transport in agricultural topsoils. *Eur J Soil Sci*. 2007;58(1):282–92.
51. Vereecken H, Kasteel R, Vanderborght J, Harter T. Upscaling hydraulic properties and soil water flow processes in heterogeneous soils: A review. *Vadose Zone J*. 2007;6(1):1–28.
52. Li J, Zhang L, Wang Y, Fredlund D. Permeability tensor and representative elementary volume of saturated cracked soil. *Can Geotech J*. 2009;46(8):928–42.
53. Abdulraheem MI, Zhang W, Li S, Moshayedi AJ, Farooque AA, Hu J. Advancement of remote sensing for soil measurements and applications: A comprehensive review. *Sustainability*. 2023;15(21):15444.
54. Babatunde C, Oyeleke O, Aruleba J. Integrating Soil Morphology, Survey Data, and Classification for Effective Land Use Planning in Tropical Regions. *Int J Plant Soil Sci*. 2025;37(8):701–12.
55. Schoeneberger PJ, Wysocki DA, Benham EC. Field book for describing and sampling soils. Government Printing Office; 2012.
56. Weber TKD, Weihermüller L, Nemes A, Bechtold M, Degré A, Diamantopoulos E, et al. Hydro-pedotransfer functions: a roadmap for future development. *Hydrol Earth Syst Sci*. 2024;28(14):3391–433.
57. Weber TKD, Weihermüller L, Nemes A, Bechtold M, Degré A, Diamantopoulos E et al. Hydro-pedotransfer functions: A roadmap for future development. *EGUsphere*. 2023.
58. McBratney AB, Santos MM, Minasny B. On digital soil mapping. *Geoderma*. 2003;117(1–2):3–52.
59. Brooks R, Corey A. Hydraulic properties of porous media. Hydrology Paper No. 3. Civil Engineering Department. Fort Collins, CO: Colorado State University; 1964.
60. Van Genuchten MT. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci Soc Am J*. 1980;44(5):892–8.
61. Mualem Y. Hysteretical models for prediction of the hydraulic conductivity of unsaturated porous media. *Water Resour Res*. 1976;12(6):1248–54.
62. Blunt MJ, Bijeljic B, Dong H, Gharbi O, Iglauer S, Mostaghimi P, et al. Pore-scale imaging and modelling. *Adv Water Resour*. 2013;51:197–216.
63. Helland J, Skjaeveland S. Relationship between capillary pressure, saturation, and interfacial area from a model of mixed-wet triangular tubes. *Water Resour Res*. 2007;43(12).

64. Beven KJ. Rainfall-runoff modelling: the primer. Wiley; 2012.
65. Šimůnek J, Van Genuchten MT, Šejna M. Recent developments and applications of the HYDRUS computer software packages. *Vadose Zone J.* 2016;15(7):vzj2016. 04.0033.
66. Steefel CI, Appelo CAJ, Arora B, Jacques D, Kalbacher T, Kolditz O, et al. Reactive transport codes for subsurface environmental simulation. *Comput GeoSci.* 2015;19(3):445–78.
67. Parkhurst DL, Appelo C. Description of input and examples for PHREEQC version 3: a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. US Geological Survey; 2013. pp. 2328–7055. Report No.
68. Lal R. Soil organic matter and water retention. *Agron J.* 2020;112(5):3265–77.
69. Parton WJ, Ojima DS, Cole CV, Schimel DS. A general model for soil organic matter dynamics: sensitivity to litter chemistry, texture and management. *Quant Model soil Form processes.* 1994;39:147–67.
70. Nearing GS, Kratzert F, Sampson AK, Pelissier CS, Klotz D, Frame JM, et al. What role does hydrological science play in the age of machine learning? *Water Resour Res.* 2021;57(3):e2020WR028091.
71. Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L. Physics-informed machine learning. *Nat Reviews Phys.* 2021;3(6):422–40.
72. Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys.* 2019;378:686–707.
73. Fuller A, Fan Z, Day C, Barlow C. Digital twin: Enabling technologies, challenges and open research. *IEEE access.* 2020;8:108952–71.
74. Ortigossa ES, Gonçalves T, Nonato LG. Explainable artificial intelligence (xai)—from theory to methods and applications. *IEEE Access.* 2024;12:80799–846.
75. M’harzi Alaoui FE, Taki Y, Aalil I, Dadah O, Ben Abbou S. From lab to algorithm: a review of AI-driven predictive modeling in soil stabilization for road construction. *J Umm Al-Qura Univ Eng Archit.* 2026:1–31.
76. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. *Adv Neural Inf Process Syst.* 2017;30.
77. Jones HG. Irrigation scheduling: advantages and pitfalls of plant-based methods. *J Exp Bot.* 2004;55(407):2427–36.
78. Lal R. Restoring soil quality to mitigate soil degradation. *Sustainability.* 2015;7(5):5875–95.
79. McBratney A, Whelan B, Ancev T, Bouma J. Future directions of precision agriculture. *Precision Agric.* 2005;6(1):7–23.
80. Dean S, Freer J, Beven K, Wade AJ, Butterfield D. Uncertainty assessment of a process-based integrated catchment model of phosphorus. *Stoch Env Res Risk Assess.* 2009;23(7):991–1010.
81. Saltelli A, Ratto M, Andres T, Campolongo F, Cariboni J, Gatelli D, et al. Global sensitivity analysis: the primer. Wiley; 2008.

82. Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans ASABE. 2007;50(3):885–900.
83. Helton JC, Johnson JD, Sallaberry CJ, Storlie CB. Survey of sampling-based methods for uncertainty and sensitivity analysis. Reliab Eng Syst Saf. 2006;91(10–11):1175–209.

Figures

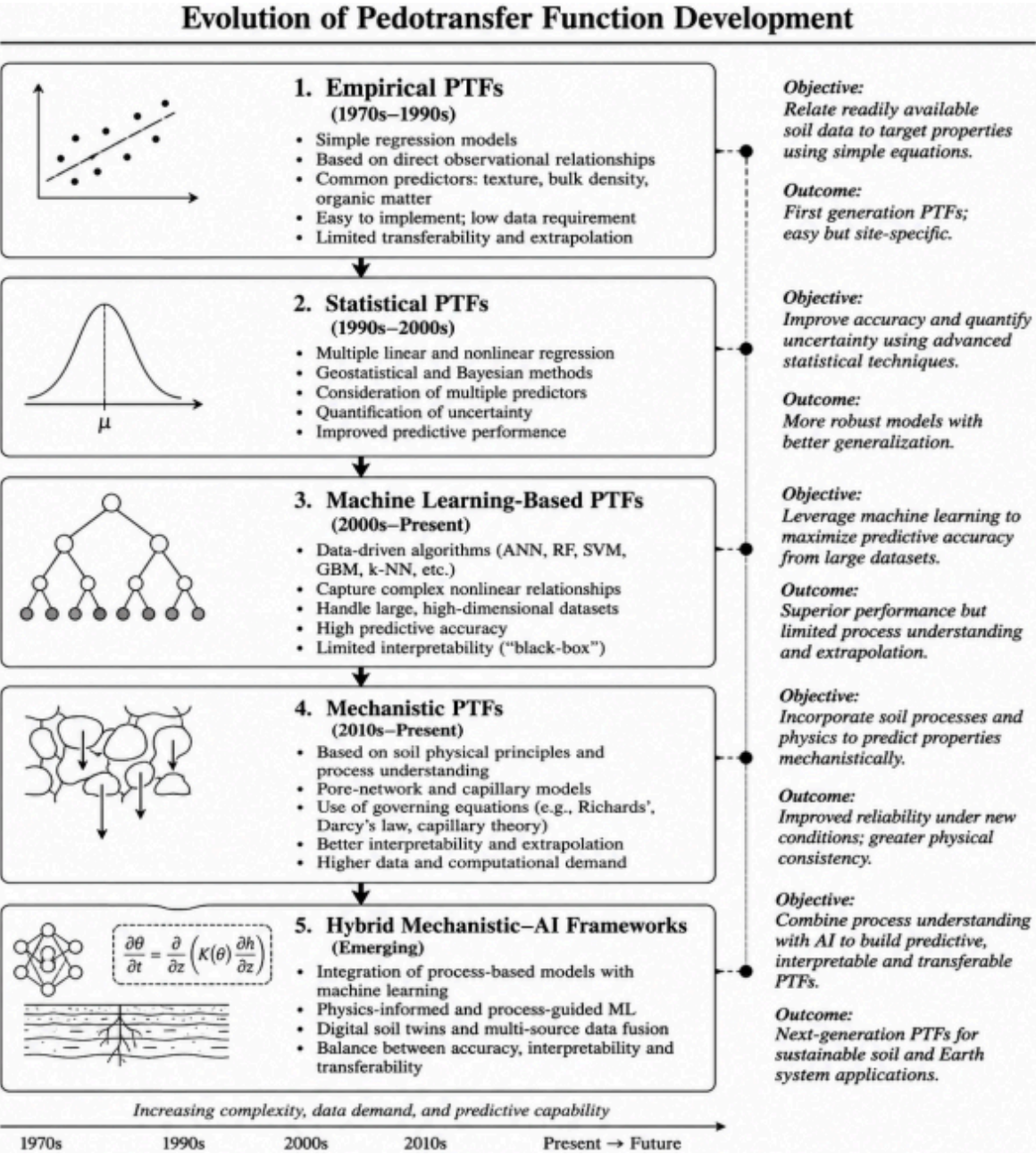


Figure 1

Evolution of pedotransfer function development from empirical to mechanistic and hybrid frameworks.

Source: Adapted from (3-6, 11, 14).

