



Informing soil models using pedotransfer functions: challenges and perspectives

Yakov Pachepsky (1) and Nunzio Romano (2)

(1) USDA-ARS Beltsville Agricultural Research Center, Environmental Microbial and Food Safety Laboratory, Beltsville, United States (yakov.pachepsky@ars.usda.gov), (2) Division of Agricultural, Forest and Biosystems Engineering, University of Naples Federico II, Naples, Italy (nunzio.romano@unina.it)

Pedotransfer functions (PTFs) are empirical relationships between parameters of soil models and more easily obtainable data on soil properties. PTFs have become an indispensable tool in modeling soil processes. As alternative methods to direct measurements, they bridge the data we have and data we need by using soil survey and monitoring data to enable modeling for real-world applications. Pedotransfer is extensively used in soil models addressing the most pressing environmental issues. The following is an attempt to provoke a discussion by listing current issues that are faced by PTF development.

1. As more intricate biogeochemical processes are being modeled, development of PTFs for parameters of those processes becomes essential.
2. Since the equations to express PTF relationships are essentially unknown, there has been a trend to employ highly nonlinear equations, e.g. neural networks, which in theory are flexible enough to simulate any dependence. This, however, comes with the penalty of large number of coefficients that are difficult to estimate reliably. A preliminary classification applied to PTF inputs and PTF development for each of the resulting groups may provide simple, transparent, and more reliable pedotransfer equations.
3. The multiplicity of models, i.e. presence of several models producing the same output variables, is commonly found in soil modeling, and is a typical feature in the PTF research field. However, PTF intercomparisons are lagging behind PTF development. This is aggravated by the fact that coefficients of PTF based on machine-learning methods are usually not reported.
4. The existence of PTFs is the result of some soil processes. Using models of those processes to generate PTFs, and more general, developing physics-based PTFs remains to be explored.
5. Estimating the variability of soil model parameters becomes increasingly important, as the newer modeling technologies such as data assimilation, ensemble modeling, and model abstraction, become progressively more popular. The variability PTFs rely on the spatio-temporal dynamics of soil variables, and that opens new sources of PTF inputs stemming from technology advances such as monitoring networks, remote and proximal sensing, and omics.
6. Burgeoning PTF development has not so far affected several persisting regional knowledge gaps. Remarkably little effort was put so far into PTF development for saline soils, calcareous and gypsiferous soils, peat soils, paddy soils, soils with well expressed shrink-swell behavior, and soils affected by freeze-thaw cycles.
7. Soils from tropical regions are quite often considered as a pseudo-entity for which a single PTF can be applied. This assumption will not be needed as more regional data will be accumulated and analyzed.
8. Other advances in regional PTFs will be possible due to presence of large databases on region-specific useful PTF inputs such as moisture equivalent, laser diffractometry data, or soil specific surface.
9. Most of flux models in soils, be it water, solutes, gas, or heat, involve parameters that are scale-dependent. Including scale dependencies in PTFs will be critical to improve PTF usability.
10. Another scale-related matter is pedotransfer for coarse-scale soil modeling, for example, in weather or climate models. Soil hydraulic parameters in these models cannot be measured and the efficiency of the pedotransfer can be evaluated only in terms of its utility. There is a pressing need to determine combinations of pedotransfer and upscaling procedures that can lead to the derivation of suitable coarse-scale soil model parameters.
11. The spatial coarse scale often assumes a coarse temporal support, and that may lead to including in PTFs other environmental variables such as topographic, weather, and management attributes.
12. Some PTF inputs are time- or space-dependent, and yet little is known whether the spatial or temporal structure of PTF outputs is properly predicted from such inputs
13. Further exploration is needed to use PTF as a source of hypotheses on and insights into relationships between soil processes and soil composition as well as between soil structure and soil functioning.

PTFs are empirical relationships and their accuracy outside the database used for the PTF development is essentially unknown. Therefore they should never be considered as an ultimate source of parameters in soil modeling. Rather they strive to provide a balance between accuracy and availability. The primary role of PTF is to assist in modeling for screening and comparative purposes, establishing ranges and/or probability distributions of model parameters, and creating realistic synthetic soil datasets and scenarios. Developing and improving PTFs will remain the mainstream way of packaging data and knowledge for applications of soil modeling.