Linking Principles of Soil Formation and Flow Regimes

Henry Lin

Dept. of Crop and Soil Sciences, The Pennsylvania State Univ., University Park, PA 16802, USA, Email henrylin@psu.edu

Abstract

Preferential flow (PF) is a fundamentally important soil hydrologic process that controls a variety of soil physical, chemical, and biological functions. However, the lack of theory in this field and the existence of conceptual and technological bottlenecks continue to hinder the advancement in understanding and predicting PF. This paper explores three theories that link pedogenesis to flow regimes, including 1) non-equilibrium thermodynamics as applied to the open dissipative system of field soils with continuous energy and mass inputs that results in the dual-partitioning of pedogenesis into organizing and dissipating processes; 2) constructal theory that explains the tendency for dual-flow regimes in soils, one with high resistivity (Darcy flow) and the other with low resistivity (PF), together they form natural PF configuration that provides the least global flow resistance; and 3) theory of evolving networks that sheds light on diverse PF networks for increasing the efficiency or effectiveness of energy and mass transfer in the subsurface. All three theories support the notion that PF is likely universal in natural soils. However, some controversies associated with these theories require more concerted efforts to systematically test their applicability and to formulate quantitative relationships between PF occurrence and its macroscopic controls.

Key Words

Preferential flow; pedogenesis; non-equilibrium thermodynamics; constructal theory; network theory.

Introduction

Preferential flow (PF) is a generic term that refers to the process whereby water (and materials carried by water) moves by preferred pathways in an accelerated speed through a fraction of a porous medium, thus bypassing a portion of the matrix. Numerous studies over the past four decades or so have demonstrated that PF can occur in practically all natural soils and landscapes. Mechanism-wise, all PFs are essentially heterogeneity-related, including macropore flow, finger flow, funnel flow, and hydrophobicity-induced flow that are more frequently reported in the soil science literature (at the pore and pedon scales), and pipe flow, return flow, throughflow, depression-focused flow, and flow at the soil-bedrock interface that are more commonly reported in the hydrology literature (at the hillslope and catchment scales).

In a new vision for watershed hydrology, McDonnell *et al.* (2007) raised a number of fundamental questions: “*Why heterogeneity exists? Why there is preferential, network-like flow at all scales?*” In light of these concerns and the need to develop PF theory, the objective of this paper is to discuss some physics-based theoretical understanding of PF occurrence and its links to soil formation and evolution.

Non-Equilibrium Thermodynamics and Dual-Partitioning of Pedogenesis

Real-world systems are not isolated from their environment and therefore are continuously exchanging energy and matter with their surroundings (including being driven by external energy sources as well as dissipating internal energy to the surroundings). It is such energy and mass flow across various gradients and boundaries that have driven the evolution and functioning of soils and ecosystems. Non-equilibrium thermodynamics can describe how soil systems interact with their surroundings and their evolution.

The second law of thermodynamics states that the entropy (S) of an isolated system not in equilibrium will tend to increase over time, approaching a maximum at equilibrium. Thermodynamic entropy can be interpreted as a measure of a system’s disorder or randomness: the higher the S, the greater the mixedupness or homogeneity (Boltzmann 1896). Interestingly, S is the *only* quantity in physical sciences that seems to imply a particular direction for time (so it is also called an arrow of time) (Prigogine 1961; Tiezzi 2003). Prigogine (1961) distinguishes two terms in the total change of entropy, dS, in an open system: the first, d_iS, is the S produced inside the system; the second, d_eS, is the transfer of S across the system boundaries. According to the second law of thermodynamics, the first term is always positive:

$$dS = d_iS + d_eS, \qquad d_iS \geq 0. \qquad [1]$$

It is in this formulation that the distinction between *reversible* and *irreversible* processes is essential (Prigogine 1961). Only irreversible processes (such as convection, diffusion, and chemical reactions) contribute to *S* production, leading to one-sidedness of time. This irreversibility results from a certain heat energy dissipation due to intermolecular friction and collisions—energy that can not be recovered if the process is reversed. From a thermodynamics perspective, all complex natural processes are essentially irreversible (Prigogine 1980), including pedogenesis and PF. Often, a threshold behaviour is involved in the time evolution of complex systems and their responses to external forcing.

In an open dissipative system like field soils, while d_iS is always positive, d_eS can be positive or negative depending on *S* exchange between the soil system and its environment (Smeck *et al.* 1983). A dissipative system is a thermodynamically open system that is operating far from equilibrium in an environment with which it exchanges energy and matter (Prigogine 1961). A dissipative system is characterized by the spontaneous appearance of *symmetry breaking* (leading to *anisotropy*) and the *formation of complex structures* (leading to *heterogeneity*) (Prigogine 1980). Heterogeneity here differs from randomness: the former is associated with order while the later is linked to disorder.

Ordering vs. Dissipative Processes in Pedogenesis

Pedogenesis is an energy-consuming process. Smeck *et al.* (1983) explained that soil systems experience outfluxes as well as influxes of energy and matter, but the net balance must favor energy influxes in order to drive soil-forming processes for soil development to proceed. Soil systems with aggregates, horizons, and profiles formed (from parent materials) over time represent more and more ordered states than their precursors (Smeck *et al.* 1983), suggesting an overall reduction in a soil system's thermodynamic *S* and an increased likelihood for PF.

Energy (*E*) inputs and *S* production result in two categories of processes during pedogenesis (Lin, 2010): (1) *Ordering processes* that lead to the formation of soil profile and soil structure, which generally involve *S* reduction, such as aggregation, humus accumulation, horizonation, flocculation, and others; and (2) *Dissipative processes* that lead to destruction of soil structure and the formation of soil matrix, which generally involve *S* increase, such as aggregate degradation, humus decomposition, erosion, dispersion, and others. Such dual-partitioning of pedogenesis results in a variety of soil architecture (which is equivalent to soil structure + soil matrix) that leads to the likelihood for PF occurrence. Similar to Prigogine's (1961) formulation, ΔS_{soil} may be partitioned as (here we use Δ instead of *d* to present longer time interval):

$$\Delta S_{soil} = \Delta S_{matrix} + \Delta S_{structure}. \qquad [2]$$

where S_{matrix} is the *S* related to dissipative processes (including *S* generated internally in the soil), while $S_{structure}$ is the *S* related to organizing processes (notably *S* exported from the soil to the surrounding). Smeck *et al.* (1983) have suggested that ΔS_{soil} for most soils are negative after grouping soil-forming processes into positive and negative ΔS: positive ΔS assigned to processes that result in disorder of soil (e.g., physical mixing and primary mineral weathering) and negative ΔS assigned to processes that sort soil constituents (e.g., leaching and accumulation of organic matter).

Near-equilibrium, Far-from-equilibrium, and Models of Soil Development

Because of the irreversible nature of pedogenesis, the second law of thermodynamics essentially dictates that all field soils will evolve towards structured heterogeneity and thus non-uniform flow. Non-equilibrium conditions may be approximated using assumptions of local equilibrium and local *S* production, which can be determined the conjugate force-flux relation.

Near-equilibrium systems tend towards a unique steady-state condition bounded by the force-flux relations. Near-equilibrium fluxes exhibit linear relations to their conjugate forces, as exemplified by the laws of Darcy, Fourier, Fick, and Ohm for water, heat, gas, and electrical transfer processes, respectively. As the forces or gradients become steeper, the linear postulates become unreliable, and systems transition through threshold and bifurcation phenomena to the nonlinear realm of far-from-equilibrium thermodynamics (Nicolis and Prigogine 1989). Open, far-from-equilibrium systems exhibit organization that is dependent on the force-flux relation and the continuous flux of energy and mass across the system boundaries. Interestingly, far-from-equilibrium systems evolve to a state of organization that most efficiently dissipates the flux of available energy, resulting in a state that maximizes the flow of energy, mass, and entropy through the system (Nicolis and Prigogine 1989; Bejan 2000).

Constructal Theory and Preferential Flow Configuration Generation and Evolution

Flow Configuration Generation and Evolution

Emerging constructal theory explains and predicts the occurrence of flow patterns in nature under a principle summarized below: *"For a finite-size flow system to persist in time (to survive), it must evolve in such a way that it provides easier and easier access to the currents that flow through it"* (Bejan 2000; 2007). This flow tendency calls for at least two flow regimes—one with high resistivity (Darcy flow) that fills the greater part of the available space and the other with low resistivity that provides fast access through various preferred pathways (such as channels and macropores). Together, the fast flowpaths and slow interstitial space constitute natural PF architecture—the configuration that offers the least global flow resistance.

While constructal theory is largely a statement without rigorous mathematical formulae at the present time and its application in the real world soils and hydrologic systems remains to be seen, this somewhat controversial theory offers an interesting perspective regarding the possible generation and evolution of PF in soils. This is because the dual-flow regime anticipated by this theory is in line with the dual-porosity system commonly reported for soils and geological materials (e.g., van Genuchten and Wierenga 1976; Flury *et al*. 1994), and it is also consistent with the pedogenic dual-partitioning discussed above.

Soil Development and Soil Hydraulic Properties Change over Time

Bejan (2000; 2007) suggested three principles to explain how flow configuration evolves over time: (1) survival by increasing flow performance; (2) survival by increasing svelteness; and (3) survival by increasing flow territory. The three constructal principles can be used to explain soil development and associated flow path generation and evolution employing. We can view flow channel openings during weathering to result from either physical breakdown of rocks or cracking, or from chemical or biological processes. Following the constructal principles, we may expect the following sequence to occur as weathering proceeds: (1) At first, as time increases, more pore space would gradually become available for flow within the weathering zone. This results in a reduced flow resistance in the weathering zone; (2) Once the above process reaches a limit (e.g., set by a maximum pore space in the material being weathered), then the weathering would expand into new parent material underneath, which will start the above 1st step again; and (3) As the overall weathering profile thickness reaches a possible limit (e.g., set by environmental constraints, such as climate), then the weathering profile would start to increase its svelteness. This would lead to the formation of a possible compacted layer in the soil profile (e.g., a degraded argillic horizon or a fragipan).

Translating the above general understanding into soil properties change over time, we may expect the following general trend of soil development (without other complicating factors): soil thickness would increase as weathering increases, but soil saturated hydraulic conductivity would first increase and then decrease in the subsoil. Furthermore, as more organization and structural heterogeneity developed through pedogenesis, more PF paths would likely to occur (either vertically or laterally or both).

Network Theory and Preferential Flow Networks in Soils

Networks Theory

A network is nothing more than a set of discrete nodes, and a set of links (representing interactions) that connect the nodes together. The nodes and their links can be anything—such as individuals (nodes) and their social interactions (links), web pages (nodes) and WWW (links), species (nodes) and food webs (links), or individual soil pores (nodes) and their connectivity (links).

In the past decade, a burst of interest in complex networks has sparked rapid growth in their theoretical studies and diverse applications, largely due to a drastic increase in the availability of network datasets coming from the Internet and electronic databases (Barabási 2009). The evolution of network theory has gone through the stages of random graph theory, small-world networks, and scale-free networks.

Network modeling has shifted from the reproduction of network's *structure* (topology) to the modeling of its *evolution* (dynamics), leading to the emerging theory of evolving networks (Albert and Barabási 2002). Another turning point in the modern view of complex networks is *preferential attachment* (Barabási and Albert 1999), meaning that new edges are not placed at random but tend to connect to vertices that already have a large degree of connectivity. Another feature in the theory of evolving networks is that processes operating at the local level both constrain and are constrained by the network structure. The inseparability of the topology and dynamics of evolving networks is increasingly recognized—though far from being fully understood. This is similar to the inseparability of soil architecture and flow dynamics in soils.

Flow Networks in Soils

Despite the difficulty of direct observations, networks are abundant in soils, such as root branching networks, mycorrhizal mycelial networks, animal borrowing networks, crack and fissure networks, and others. Energy inputs cause flow networks to form in soils, and networks provide a means of minimizing energy dissipation. Like the energy of water flowing over the land surface that creates dendritic stream networks, water flowing through soils also creates network-like flow paths in the subsurface. As water moves through soils, changes in soil texture, structure, organic content, mineral species, biological activities, and other features will modify the resistance to the flow, causing change in flow path to allow water to follow the least resistant path, thus resulting in a PF network that has the least global flow resistance. Some evidence has suggested that a similarity may exist between river dendritic structure and subsurface PF networks.

Depending on the governing physical processes, flow networks may exhibit different topologies. Overall, flow and transport networks in soils are formed by the forcing of soil formation, mainly climate and organisms. Cycles of wetting and drying, freezing and thawing, shrinking and swelling, coupled with organic matter accumulation and decomposition, biological activities, and chemical reactions have led to the formation of diverse soil aggregates and pore networks in the subsurface. In particular, plant roots, burrowing animals, and mycorrhizal mycelia are active in creating networks in soils. Common PF networks in soils include crack and interpedal networks, root networks, mycorrhizal mycelial networks, animal borrowing networks, and man-made subsurface drainage networks.

Various vertical and lateral PFs in soils constitute subsurface flow networks that dictate how water percolates through the soil, runs down the hillslope, and moves across the watershed. The origin, dynamics, and recurrent patterns of self-organization of such flow networks in the subsurface have become the subjects of recent research and model development. For instance, during storms with wet soil conditions, a subsurface network often provides preferred pathways for water to flow down gradient with high velocities. Thresholds may occur when significant changes happen rapidly. A number of studies have suggested that precipitation thresholds for subsurface stormwater generation may be widespread phenomena. Even individual short PF pathways can be linked via a series of nodes in a network, which may be switched on or off and expand or shrink depending on local soil moisture conditions, rainfall inputs, or landscape positions.

Conclusion

The lack of PF theory in field soils requires concerted efforts to synthesize concepts and to advance techniques for measuring and modeling PF across space and time. The theories of non-equilibrium thermodynamics, constructal theory, and evolving networks provide some encouraging perspectives towards a physics-based understanding and prediction of PF that is closely linked to soil formation and evolution.

References

Albert R, Barabási AL (2002) Statistical mechanics of complex networks. *Reviews of Modern Physics* **74**, 47-97

Barabási AL (2009) Scale-free networks: A decade and beyond. *Science* **325**, 412-413..

Barabási AL, Albert R (1999) Emergence of Scaling in Random Networks. *Science* **286**, 509-512.

Bejan A (2000) Shape and Structure, from Engineering to Nature. Cambridge Univ. Press, Cambridge, UK.

Bejan A (2007) Constructal theory of pattern formation. *Hydrology and Earth System Sciences* **11**, 753-768.

Boltzmann L (1896) Lectures on Gas Theory. Dover Publications (March 27, 1995).

Flury M, Flühler H, Jury WA, Leuenberger J (1994) Susceptibility of soils to preferential flow of water: a field study. *Water Resources Research* **30**, 1945–1954.

Lin, H.S. 2010. Comments on Energy-based Pedogenic Models by Field and Minasny (2008) and Rasmussen (2008). *Soil Science Society of America Journal* **74**, 1-3.

McDonnell JJ, Sivapalan M, Vache´ K, Dunn S, Grant G, Haggerty R, Hinz C, Hooper R, Kirchner J, Roderick ML, Selker J, Weiler M (2007) Moving beyond heterogeneity and process complexity: A new vision for watershed hydrology, *Water Resources Research* **43**, W07301.

Nicolis G, Prigogine I (1989) Exploring Complexity. W. H. Freeman, New York.

Prigogine I (1961) Introduction to thermodynamics of irreversible processes. John Wiley, New York, NY.

Prigogine I (1980) From being to becoming : time and complexity in the physical sciences. W. H. Freeman.

Smeck NE, Runge ECA, Mackintosh EE (1983) Dynamics and genetic modeling of soil systems. In ‘Pedogenesis and Soil Taxonomy’. (Eds LP Wilding *et al.*) pp. 51–81. (Elsevier, New York).

Tiezzi E (2003) The essence of time. WIT Press.

van Genuchten MT, Wierenga PJ (1976) Mass transfer studies in sorbing porous media Part 1 analytical solutions. *Soil Science Society of America Journal* **40**, 473-480.