Modeling Soil Processes: Review, Key Challenges, and New Perspectives


The remarkable complexity of soil and its importance to a wide range of ecosystem services presents major challenges to the modeling of soil processes. Although major progress in soil models has occurred in the last decades, models of soil processes remain disjointed between disciplines or ecosystem services, with considerable uncertainty remaining in the quality of predictions and several challenges that remain yet to be addressed. First, there is a need to improve exchange of knowledge and experience among the different disciplines in soil science and to reach out to other Earth science communities. Second, the community needs to develop a new generation of soil models based on a systemic approach comprising relevant physical, chemical, and biological processes to address critical knowledge gaps in our understanding of soil processes and their interactions. Overcoming these challenges will facilitate exchanges between soil modeling and climate, plant, and social science modeling communities. It will allow us to contribute to preserve and improve our assessment of ecosystem services and advance our understanding of climate-change feedback mechanisms, among others, thereby facilitating and strengthening communication among scientific disciplines and society. We review the role of modeling soil processes in quantifying key soil processes that shape ecosystem services, with a focus on provisioning and regulating services. We then identify key challenges in modeling soil processes, including the systematic incorporation of heterogeneity and uncertainty, the integration of data and models, and strategies for effective integration of knowledge on physical, chemical, and biological soil processes. We discuss how the soil modeling community could best interface with modern modeling activities in other disciplines, such as climate, ecology, and plant research, and how to weave novel observation and measurement techniques into soil models. We propose the establishment of an international soil modeling consortium to coherently advance soil modeling activities and foster communication with other Earth science disciplines. Such a consortium should promote soil modeling platforms and data repository for model development, calibration and intercomparison essential for addressing contemporary challenges.

Abbreviations: AIC, Akaike information criterion; BE, Bayesian model evidence; BIC, Bayesian information criterion; BMA, Bayesian model averaging; CLM, Community Land Model; DEM, digital elevation model; EnKF, Ensemble Kalman Filter; ET, evapotranspiration; GHG, greenhouse gases; GIS, geographic information system; GPS, global positioning system; IC, information criterion; ISMC, International Soil Modeling Consortium; KIC, Kashyap information criterion; LIDAR, Light Detection and Ranging; MCMC, Markov chain Monte Carlo; MRT, magnetic resonance imaging; MW, microwave spectrum; MWIR, mid-wave infrared spectrum; NIR, near-infrared spectrum; OTU, operational taxonomic units; pdf, probability density function; PSS, proximal soil sensing; PTF, pedotransfer function; SAR, Synthetic Aperture Radar; SDA, sequential data assimilation; SVAT, soil–vegetation–atmosphere transfer; SWIR, short-wave infrared spectrum; TE, treated effluents; TIR, thermal infrared spectrum; UAV, unmanned air vehicles; µCT, microcomputed tomography; VIS, visible spectrum; VSP, virtual soil platform.

*Corresponding author (h.vereeken@fz-juelich.de). See end of paper for author affiliations.

Vadose Zone J. doi:10.2136/vzj2015.09.0131
Received 28 Sept. 2015. Accepted 29 Mar. 2016. Open access
Vol. 15, Iss. 5, 2016
© Soil Science Society of America.
This is an open access article distributed under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
A Brief History of Soil Modeling

The quantitative description of physical, chemical, and biological interactions in soil at multiple scales and levels of refinement has been a long-standing goal and key challenge in soil science. The earliest numerical and analytical models in the field of soil science date back to the last century and dealt mainly with the simulation of water flow (e.g., Hanks and Bowers, 1962; Rubin and Steinhardt, 1963; Whisler and Kulte, 1965; Bresler and Hanks, 1969; Van Keulen and Van Beek, 1971), heat flow (Wierenga and De Wit, 1970), solute transport processes (Dutt and Tanji, 1962; Lindstrom et al., 1967; Bear, 1972; Bresler, 1973; Gerke and van Genuchten, 1993), soil organic carbon dynamics (Russell, 1964, 1975; Van Veen and Paul, 1981), and nutrient dynamics (Kirkham and Bartholomew, 1955; Cole et al., 1978). These models consisted mostly of analytical solutions of partial differential equations for well-defined soils and porous media, numerical solutions of single partial differential equations, or conceptual models that were solved with analog or digital computers.

These first generation models that proliferated with the availability of the digital computer focused primarily on physical and chemical processes without explicit consideration of biotic processes or accounting for the role of soil structural related processes. One of the first models addressing the role of soil structure in the decomposition of organic matter by microorganisms was developed by Van Veen and Paul (1981) and Van Veen et al. (1985) and reviewed in Van Veen and Kuikman (1990). An early model that considered the role of soil structure on solute transport and leaching was developed by Addiscott (1977). The role of soil structure on soil physical processes, including water flow and solute transport, was conceptualized and framed in a mathematically consistent approach in the early 1990s by Gerke and van Genuchten (1993). A first suite of soil ecosystem dynamics models including detrital food webs was published in the early 1970s by Patten (1972) and McBryar et al. (1977), and in the 1980s by Rosswall and Paustian (1984) and later de Ruiter et al. (1993). These studies addressed the role of soil microbes and soil fauna within the framework of food webs and nutrient dynamics. Recently, soil ecosystem models have been developed that allow modeling of soil biodiversity and its loss, as well as the role of microbes and soil fauna in soil nutrient transfer processes (Hunt and Wall, 2002).

Due to availability of novel measurement and analytical techniques such as X-ray tomography, soil neutron tomography, magnetic resonance imaging, as well as molecular techniques that enable quantification of molecular-driven soil biological processes and soil microbial composition, data have now become available that allow the development and validation of soil models that are able to quantify physical, chemical, and biological processes at the level of the pore scale and smaller. Combined with an increased understanding of the complex interactions of soil processes, the advent of computers and progress in the development of analytical and improved numerical algorithms, especially at the end of the 1980s, have empowered the development of complex soil models integrating physical, chemical, and biological processes from the pore scale to the global scale (Parton et al., 1998). Notwithstanding the considerable progress from early modeling efforts, fundamental soil processes and their interactions remain lacking and deficient, such that it hampers the prediction and quantification of key soil functions and services. Moreover, the integration and quantification of available knowledge on soil processes remain sketchy due to lack of coherence and limited communication among research communities and disciplines.

In this paper we review the state-of-the-art soil modeling as well as...
The State-of-the-Art of Modeling Soil Processes

Advanced soil models today use the Richards equation and the convection–dispersion equation to describe water and solute movement through soils, and often are able to account for preferential flow and transport (Šimůnek et al., 2003). Many of these models include the simulation of heat flow and energy balance approaches, providing information on soil temperature dynamics and water vapor flow. Soil chemistry ranges from simple equilibrium or nonequilibrium sorption models to complex multi-species models, such as that by Jacques et al. (2008). For contaminated soils, the typical single phase flow models have been extended to include multiphase flow phenomena to take into account complex interactions between solid, liquid, gas, and contaminant phases. Soil carbon dynamics are typically conceptualized by multi-compartment approaches, where each compartment is composed of organic matter with similar chemical composition or degradability (Coleman et al., 1997; Bricklemyer et al., 2007). Nitrogen turnover is strongly related to carbon turnover, and both are often part of an overall model of C, N, and nutrient cycling in terrestrial ecosystems (Priesack et al., 2008; Manzoni and Porporato, 2009; Batlle-Aguilar et al., 2011). Compared to the above process descriptions, several process descriptions presented below are still in their infancy. At present, many soil models consider the soil to be a rigid medium. Yet, we know that management practices and natural events such as droughts and floods may drastically change soil’s architecture and structure. The description of root water uptake is mostly based on simple approaches such as the model of Feddes et al. (1976). Only recently more complex approaches that explicitly describe the three-dimensional soil root system have become available (Hopmans and Bristow, 2002; Schroder et al., 2008; Javaux et al., 2013) and are not yet widespread. Improved descriptions of root solute uptake include root hairs, root exudation, and rhizodeposition, which increases microbial activity (Kuzyakov and Domanski, 2000), or the role of arbuscular mycorrhizal fungi (Schnepf et al., 2008b; Leitner et al., 2010; Schnepf et al., 2012). However, these improved descriptions are not yet sufficiently incorporated into soil–crop models (Hinsinger et al., 2011). There is an overall lack of spatially explicit models that properly describe soil C and nutrient dynamics at different spatial scales (Manzoni and Porporato, 2009). Approaches to simulating temporal changes of soil structure, a major determinant of water movement, biological activity and root growth, and soil erosion, are relatively rare and at an early stage of development (Leij et al., 2002; Stamati et al., 2013). There are few models of interactions between physical and biological processes (Tartakovsky et al., 2009; Laudone et al., 2011). However, the impact of soil biodiversity on soil productivity, crop growth, and yield has hardly been included in current soil simulation models. Recent advances in measurement technologies have provided new insights about the role of soil biodiversity on soil and crop processes, generating new knowledge and opening new perspectives for their mathematical description.

The Role of Soil Modeling in Quantifying Ecosystem Services

We capitalize on the framework of ecosystem services to analyze challenges and offer perspectives on soil modeling. Soil plays a prominent role in regulating and provisioning ecosystem services, as well as degrading and supporting processes, all linked to societal and population issues and central to scientific underpinning of how the planet functions (Adhikari and Hartemink, 2016). We rely on the conceptual framework of Dominati et al. (2010) to frame soil modeling activities related to the description and prediction of soil processes and properties (Fig. 1). The Dominati framework offers a holistic view of how soil processes and related ecosystem services are impacted by external drivers (both natural and anthropogenic) and affecting processes and soil natural capital. The various components and subcomponents, including basic processes, natural capital of soils, and ecosystem services, can be harnessed to meet human needs. But, these can also be impacted by changes in land use, agricultural practices, technological developments, climate change, and natural hazards. The natural capital of soils is defined as the stocks of mass and energy in the soil and their organization (entropy) (Robinson et al., 2009, 2014). It is related to the notion of soil properties, some of which are considered inherent and others which can be modified through management. The paper addresses a range of soil modeling activities that attempt...
Regulating services provide means to humans to live in a stable, healthy and resilient environment (Dominati et al., 2010). They include climate regulation, water regulation, erosion control, buffering, and filtering. Climate regulation is defined as the capacity of the soil to control states and fluxes of energy, water and matter that impact climate. Water regulation comprises services of the soil related to storage and retention of quantities of water. It impacts soil hydrological processes, such as runoff, leaching, and ground-water recharge, and water management practices, such as irrigation and drainage. Soils have the capacity to store and release chemicals, thereby controlling soil, water, crop, and air quality. Provisioning services are related to products derived from ecosystems (e.g., food, wood, fiber, fresh water, physical support, and genetic resources), in all of which soils play a key role. Underlying these processes are basic biological, physical, and geochemical processes. Most soil modeling research thus far has been focused on addressing these basic processes independently or coupled with a limited set of basic processes. The goal of this review is to present the key roles of state-of-the-art soil modeling approaches. The key questions addressed here are how soil modeling activities can better serve quantification of soil processes and related ecosystem services, and what areas and key challenges need to be addressed to improve the applicability and usefulness of current soil models. We substantially expand on the review paper by Jury et al. (2011), which mainly focused on the status and challenges in soil physics research dealing with soil physical methods and approaches to characterize soil water properties, scaling and effective hydraulic properties, unstable flow and water repellency, effects of plants on transport processes, characterizing soil microbial diversity, and the role of soil ecology in providing ecosystem services.

In the field of ecosystem services research, mechanistic descriptions of soil functions used to quantify ecosystem services are rarely used. Typical approaches for quantification of ecosystem services comprise the use of one-dimensional proxies based on land use or land cover, nonvalidated models based on likely combinations of explanatory variables derived from expert knowledge, validated equations that are calibrated on primary and secondary data, representative data collected within the area used to quantify ecosystem services, and implicit modeling of the ecosystem service quantity within a monetary value transfer function (Schägner et al., 2013). All these approaches aim at quantifying the supply side of ecosystem services, and several models have been developed in this respect, such as the polyscale model (Jackson et al., 2013), Invest (Nelson et al., 2009), ARIES (Villa et al., 2014), and SolVES (Sherrouse et al., 2011). However, testing and validating the accuracy and precision of these models and approaches are still open issues that need to be addressed (Schägner et al., 2013). Ensemble calculations of ecosystem services using different model approaches including more complex mechanistic models to quantify specific ecosystem functions may be valuable in quantifying uncertainty following approaches similar to the ensemble calculations used in IPCC. The proposed modeling and intercomparison platform presented later in this paper may provide an excellent opportunity to perform these kinds of analyses. Using more complex models based on a mechanistic representation of soil processes may serve as benchmarks for selecting the simpler models in these ecosystem service models. In addition, highly complex modeling approaches have the potential to be simplified to be more easily embedded in such models. One of the main reasons for using simplified models vs. more complex models for the assessment and quantification of soil processes is the issue of data availability. Soil science research has developed several approaches to address this data scarcity issue, such as proximal soil sensing, pedotransfer functions, and remote sensing of the soil surface. These approaches have not yet been applied in ecosystem services quantification. Also, issues of spatial variability of key ecosystem properties, a topic that has been at the core of soil science research, has been identified in ecosystem services research as a topic warranting more attention to better assess uncertainty (Schägner et al., 2013). In recent years, the spatial mapping of ecosystem services has gained increasing importance, while in combination with geographic information system (GIS) methods larger areas can be quantified, and spatial patterns of ecosystem services can be better identified (e.g., Calzolari et al., 2016). In the mapping of ecosystem services, such a GIS framework typically constitutes the core engine of ecosystem service quantification, allowing the quantification not only of vertically driven and local scale processes but also laterally controlled processes, such as overland flow, routing and erosion, and sediment transport (e.g., polyscale model). Also, there is a need to assess the quality and accuracy of the predictions and to validate them against real world data.

The demand for ecosystem services is mainly determined by socioeconomic characteristics and drivers, and its quantification or valuation has mainly been addressed by economists. There are several approaches used in ecosystem services research to do this valuation, but this is outside the scope of this paper. We refer the reader to the work of Bateman et al. (2013), Schägner et al. (2013), and Obst (2015), among others. Here we address soil mechanistic models and their application to quantify ecosystem services. Figure 1 illustrates the link between soil processes, soil natural capital, and ecosystem services from a soil modeling perspective (adapted from Dominati et al., 2010).
Table 2 shows in an exemplary manner a number of published studies in which mechanistic soil models have been used to quantify ecosystem services. A community-supported list of soil models can be found in Aitkenhead (2016) and at https://soil-modeling.org/models (accessed 26 Apr. 2016).

Both Fig. 1 and Table 2 are organized along the ecosystem services provided by soil, and the entries in Table 2 refer to the ecosystem services presented in the following two major sections, “Modeling Soil Supporting and Degrading Processes” and “Soil Modeling and Ecosystem Services.” Both sections are organized along these entries.

**Modeling Soil Supporting and Degrading Processes**

In this section, we will address the state of modeling soil processes with respect to quantifying soil supporting and degrading processes. They directly influence soil structure, architecture, and basic soil properties, thereby affecting the regulating and provisioning services. As shown in Fig. 1, supporting processes include the formation of soil, cycling of water and nutrients, and biological activity. Degrading processes include salinization, erosion, and compaction. At the end of the section we present five key challenges to modeling soil supporting and degrading processes (Table 3).

**Supporting Processes**

**Soil Formation**

Soil formation refers to the combination of physical, chemical, biological, and anthropogenic processes acting on a soil parent material over periods from years to millennia. Human activities, often related to agricultural practices, strongly contribute to short-term soil formation by causing aggregation, compaction, leaching, clay migration, salinization, and changes in the carbon stock. Many specific modeling studies focus on leaching (Dann et al., 2006; Jabro et al., 2006), soil carbon change (Smith et al., 1997), soil acidification (Kros et al., 1999), compaction (Nawaz et al., 2013), or other processes. However, few models treat soil formation as a coevolution of a large number of soil parameters (Finke and Hutson, 2008) in an integrated approach, thus limiting pedogenetic modeling progress (Opolot and Finke, 2015).

Soil formation is often associated with volumetric changes from strain (Brimhall and Dietrich, 1987) because of elastic and inelastic responses to pressure, decalcification, clay transport, and perturbations of different types, including tillage and bioturbation. However, most models assume a constant soil volume, neglecting changes in macroporosity and the dynamic impact of changing water quality on soil hydraulic properties. Thus, most soil models ignore soil structure dynamics and its relevance to the physical isolation of soil components like soil organic carbon by aggregation (Six et al., 2002; Six and Paustian, 2014). This may seem insignificant for short-term studies, but changes in soil structure are key processes at the time scales of decades and centuries associated with long-term soil formation. For example, short time scale processes of colloid transport are key in pedogenetic clay migration (illuviation) in soil profile development.

Volume strain also induces soil heterogeneity, as both aggregation and compaction affect macroporosity and may cause high spatial variability in surface and subsurface flow and transport processes, and in turn affect local rates of soil erosion and soil formation. For example, preferential flow may cause persistent leaching pathways at short (leaching hot spots; Koestel et al., 2013), and long time scales (persistent leaching through ripening cracks and albeluviic tongues; Sauer et al., 2009). Research questions remain on development of soil heterogeneity with time, and the possible self-enforcing or self-limiting mechanisms, as well as the relevant spatial scales with appropriate upscaling and downscaling techniques (Bierkens and de Willigen, 2000). At pedogenetic time scales, boundary input values are uncertain, meaning that climate, vegetation, and historic human activities are highly uncertain as well, and influence the degree to which soil models can be calibrated. The effect of such uncertainties must be determined to allow for accurate scenario-like quantification of ecosystem services under global change.

**Water Cycling**

Water cycling in soils involves the infiltration of precipitation in soils and the subsequent release of this water to the atmosphere, and groundwater and surface water systems by evapotranspiration (ET) and leaching, respectively. To characterize and predict ecosystem services provided by soils, we must quantify the amount of rainfall, interception, soil infiltration, soil moisture redistribution and root water uptake. Among these processes, rainfall is highly variable in space and time, difficult to measure and extremely difficult to predict (Villarini, 2009). In addition, climate change will lead to an increase in its spatial-temporal variability and intensity (e.g., strong convective rainfall events) challenging the quantification of infiltration and overland flow processes. For soil moisture redistribution, common soil water flow models employ the Richards equation, which combines the Darcy equation with the continuity equation; including a sink term for soil water extraction by roots (see Eq. [1]).

$$\frac{\partial \theta}{\partial t} = \nabla \cdot (K \nabla H) - S$$  \[1\]

where $\theta$ is the volumetric water content ($L^3 L^{-3}$), $t$ is the time ($T$), $K$ is the unsaturated soil hydraulic conductivity tensor ($L T^{-1}$), $H$ is the hydraulic head ($L$), and $S$ is the sink term accounting for root water uptake ($L^3 L^{-3} T^{-1}$). A description of these basic processes and methods to solve this equation were provided by Aksoy and Kavvas (2005), Feddes et al. (1988). Some of the frequently used
Table 2. Examples of studies in which soil models have been used to quantify ecosystem services.

<table>
<thead>
<tr>
<th>Ecosystem services and soil processes</th>
<th>Numerical package to model this ecosystem service or soil process</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Supporting processes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Soil formation</td>
<td>SoilGen2, Soil-Landscape Model, Orthod Model, mARM3D, MILESD</td>
<td>Finke (2012), Finke and Hutson (2008)</td>
</tr>
<tr>
<td>Water cycling</td>
<td>HYDRUS-1D, WaSim-ETH, Community Land Model (CLM), SiSPAT-Isotope, SWAP</td>
<td>Karimov et al. (2014), Krause and Bronstert (2007)</td>
</tr>
<tr>
<td>Nutrient cycling</td>
<td>HPI, RothC, Century, SWAP, SiSPAT-Isootope, SWAP, S ( \Delta ) P</td>
<td>Thaysen et al. (2014), Dungait et al. (2012), Parton et al. (1993), Perego et al. (2012), Bonfante et al. (2010)</td>
</tr>
<tr>
<td>Biological activity</td>
<td>DEMENT, Soil food web model</td>
<td>Allison (2012, 2014), Holtkamp et al. (2011)</td>
</tr>
<tr>
<td><strong>Degrading processes</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Salinization</td>
<td>HYDRUS-2D, UNSATCHEM, SALTMED, SODIC, SWAP</td>
<td>Ramos et al. (2012), Schoups et al. (2006), Ragab (2002)</td>
</tr>
<tr>
<td>Erosion</td>
<td>KINEROS, WEPP, SIDISS</td>
<td>Nedkov and Burkhard (2012), Savabi et al. (1995), Simota et al. (2005)</td>
</tr>
<tr>
<td><strong>Regulating services</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Climate regulation</td>
<td>CLM4.5, SWAP</td>
<td>Oleson et al. (2013), Pollacco and Mohanty (2012)</td>
</tr>
<tr>
<td>Buffering and filtering</td>
<td>HPI, SWAP</td>
<td>Letterme et al. (2014), Bonten et al. (2012)</td>
</tr>
<tr>
<td><strong>Provisioning services</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical support</td>
<td>Volumetric Soil Model, Slope Stability Model, BBA, RipRoot, Self_org</td>
<td>Sheng et al. (2008), Arrouays et al. (2014), Alonso et al. (1990), Pollen-Bankhead and Simon (2009), Crawford et al. (2011)</td>
</tr>
<tr>
<td>Soil and habitat</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The spatial and temporal dynamics of soil water flow are usually assumed to be controlled by the soil’s unsaturated hydraulic conductivity and hydraulic head gradients; in cultivated topsoils, both vary rapidly in space and time. Soils with shallow groundwater levels may show a continuous alternation of percolation and capillary rise (Li et al., 2014). Soil heterogeneity is caused by both soil deposition and formation, as well as by land use and soil management practices. Pore-scale models have been developed to generate the change of soil hydraulic properties due to compaction, shearing and shrinkage (Alaoui et al., 2011). Soil heterogeneity may cause preferential flow through macropores and flow instabilities (Šimůnek et al., 2003), which will reduce soil water residence time and accelerated soil chemical transport. Also, hydrophobicity and wettability of soil surfaces may induce preferential flow processes and nonhomogeneous movement of water in soils (Dekker and Ritsema, 1994) (Ritsema et al., 1993). Despite being in use for more than a century, Richards equation–based models are still not suitable for all soil types (particularly soils with high clay or organic matter contents), and there is still not an adequate physical theory linking all types of flow (Beven and Germann, 2013).

Soil water and root zone processes are fundamental to the well-being of plants because they control the transport of nutrients and assimilates from photosynthesis, facilitate numerous chemical reactions, and indirectly support the transport of hormones, cell turgor, and cooling of leaves by transpiration due to root water uptake (Ehlers and Goss, 2003). Soil water flow and vegetation development are therefore closely related. For example, in periods with low leaf area index, rainfall interception and root water uptake are reduced, which may enhance runoff. Vice versa, soil moisture and oxygen availability have a large influence on vegetation growth. Existing agrohydrological models typically focus on the soil’s physical processes and treat transpiration, root water uptake, and crop development in a simplified way. In contrast, common crop and agronomic models include detailed CO₂ assimilation and plant organ development modules, but lack rigorous description of soil root zone processes. To address the close interactions between vegetation and soil, future models must better integrate soil physical knowledge with agronomic and plant physiological knowledge. Main challenges include the simulation of root development and soil water uptake, plant transpiration and vegetation growth in response to heterogeneous soil conditions. Crop root water extraction should account for root density, soil hydraulic functions, root mucilage, soil water status, and the suction of roots. Regarding crop transpiration, coupled crop–soil models should apply Penman–Monteith without the empirical crop factor (Shuttleworth, 2006). Typically, crop coefficients are being used to adapt the predicted reference ET for a well-watered grass cover to the specific crop (Farahani et al., 2007). The stomatal resistance plays a key role in transpiration, and it is controlled by solar radiation, air temperature, air humidity, CO₂ concentration, and leaf water potential. In addition, leaf area index, plant height, albedo, and nonuniform soil moisture distribution need to be accounted for (Kool et al., 2014). A large number of initiatives to integrate soil water flow and plant growth exist (Romano et al., 2011; van Lier et al., 2013; Wöhlting et al., 2013; Gayler et al., 2014). To better address the water cycle at a range of scales, there is a need for more efficient integrated modeling tools, which will be elaborated in the final major section, “Toward a Soil Modeling Platform.” The models described in this section are based on the assumption that the soil is a rigid porous medium. Soil structural dynamics will be discussed in the “Soil Formation” section and have been addressed by, for example, Or et al. (2000), and Basu and Kumar (2014).

### Nutrient Cycling

The availability of plant nutrient elements often limits plant productivity in natural and agricultural ecosystems (Marschner and Marschner, 1995). Since primary production is strongly linked to provisioning services and carbon sequestration and is often inversely related to biodiversity, the cycling of nutrients is a supporting process that has strong effects on ecosystem services (e.g., see “Biomass Production for Food, Fiber, and Energy” section). In natural systems, nutrient inputs from weathering and deposition are generally very
limited, and biomass and soil C stocks are governed by long-term rates of influx and loss. In agriculture and production forestry, productivity is often boosted by fertilizer and manure additions, but the cycling of nutrients remains important in determining nutrient use efficiency, the maintenance of nutrient stocks, and groundwater pollution. Management has major effects on nutrient cycling.

Nutrient transport in soil is intrinsically linked to water flow (see “Water Cycling” and “Linking Soil Modeling Platforms with Climate, Ecology, and Hydrology” sections). Most soils receive a net throughput of water at least in certain seasons. This is important for preventing salinization, but it also means that plant nutrients can easily be leached beyond the rooting zone, particularly during the early stages of crop growth (Rowe et al., 2001). The main aim of predictive models of nutrient cycling is to quantify the availability in time and space of nutrient elements in soil and to assess likely effects on plant growth and on nutrient loss fluxes, which can affect water and air quality. Quantifying nutrient availability requires an understanding of the rates with which nutrient elements enter, move within, and leave the soil and are mineralized from organic materials (Havlin et al., 2013). Transport and leaching of nutrients and other dissolved substances in soils are typically described by the convection–dispersion equation (CDE):

$$\frac{\partial ([\theta + \rho s]e)}{\partial t} = \nabla (\theta D_e \nabla c - q e) - S_t$$  \[2\]

where $\theta$ is the soil moisture content (L$^3$ L$^{-3}$), $c$ is the concentration of a substance in the liquid phase (M L$^{-3}$), $s$ is the concentration of the component in the solid phase, $D_e$ is the effective dispersion tensor (L$^2$ T$^{-1}$), $q$ is the Darcy flux of water (L T$^{-1}$), which is typically obtained from solving the Richards equation (Eq. [1]), $S_t$ is the sink term for nutrient uptake by roots (M L$^{-3}$ T$^{-1}$). For linear equilibrium sorption, the left term of Eq. [2] becomes:

$$\frac{\partial ([\theta + \rho K_d]e)}{\partial t}$$

where $K_d$ is the distribution coefficient (L$^3$ M$^{-1}$).

Nutrient cycling models must take into account the major fluxes of nutrient elements into soil via litter, animal excreta, and manures and fertilizers and already predict nutrient availability fairly well, particularly in response to mineral fertilizers (e.g., Archontoulis et al., 2014). More difficult to predict are microbial-mediated fluxes, such as organic nutrient mineralization rates that can be enormously variable. Predictions of mineralization rates of organic materials have frequently been based on their composition in terms of element stoichiometry, on compounds that are relatively labile or recalcitrant, and/or compounds that directly inhibit enzyme activity such as soluble phenolics. Plants also exert strong control on the soil nutrient system, indirectly by determining nutrient and carbon inputs in litter, but also directly by depleting solutes, and by accelerating removal of nutrients from minerals and organic matter mineralization via exudates, exo-enzymes and mycorrhizae. Nutrient cycling models are increasingly taking these effects into account (Taylor et al., 2011).

The mineralization and transformation of plant litter and soil organic matter has mainly been modeled using schemas of conceptual pools that turn over at different rates; these were reviewed recently by Manzoni and Porporato (2009) and Falloon and Smith (2010). For example, the Roth-C model (Coleman et al., 1997) splits litter into “resistant” and “decomposable” material and soil organic matter into “microbial,” “humified,” and “inert” material and tracks transfers among these pools using first-order rate coefficients. The values of these coefficients are modified according to temperature, moisture, and soil cover. Similar schemas are used in CENTURY (Parton et al., 1988), DAISY (Hansen et al., 1991), and ECOSSE (Smith et al., 2010), among other models. Several challenges exist with this approach. Most turnover is of recent material, but the bulk of the organic matter in soil is relatively old. Understanding how nutrients will be incorporated into and released from this large stock depends on quantifying transfers into more inert pools, which are relatively small and difficult to observe. Given several organic matter pools and unconstrained rate coefficients, it is possible to reproduce a very wide range of decomposition trajectories, which limits the predictive ability of these models. Predictions of nutrient cycling rates are likely to be improved by constraining models using actual measurements of element stocks and fluxes. The average age of soil organic carbon obtained through $^{14}$C dating is a particularly useful measurement, and is used in models such as N14C (Tipping et al., 2012) (Fig. 2) to reduce the number of unconstrained parameters. An additional way forward in flux quantification is stable isotope tracking;
As well as providing nutrient inputs in litter, plants influence nutrient cycling by removing nutrient elements from the soil solution as they become available either in mineral form or as small organic molecules (Chapin et al., 1993). The efficiency of this process means that observed nutrient concentrations in soil solution are often close to zero during active plant growth. A major challenge in modeling nutrient availability is therefore determining the most appropriate measurement with which to compare model predictions (Schimel and Bennett, 2004). Time-integrated measurements, such as net mineralization (Rowe et al., 2011) or sorption onto resins (Qian and Schoenau, 2002), are generally preferable. The prediction of nutrient availability in terms of a metric that is measurable remains a key goal for soil nutrient modeling.

Considerable progress has been made to resolve rhizospheric processes (see “Biomass Production for Food, Fiber, and Energy” section), yet mechanistic modeling of the direct effects of plants on nutrient release from organic matter and weatherable minerals through root exudation and enzyme production are currently limited to a few models of nutrient cycling at the ecosystem scale. Organic acids exuded by roots or microbes can increase nutrient solubility via effects on the pH of microsites, and/or provide a source of labile C, which allows bacteria and fungi to mineralize more recalcitrant substrates. Accounting for root exudates is important because comparatively small exudate fluxes can have a disproportionate effect in increasing nutrient availability (Yin et al., 2014). Roots and mycorrhizae also produce enzymes that directly solubilize nutrients. Production of such enzymes may be limited by N availability, sometimes leading to counterintuitive responses, such as increasing plant tissue P content with increasing N inputs (Rowe et al., 2008).

Many studies of nutrient cycling have addressed only a single element, most commonly N. Nitrogen is the nutrient element required in largest quantities, but the cycling of N into and out of plants can be controlled by other elements. Productivity in natural systems may ultimately be limited by the availability of elements essential for N fixation, such as P or Mo (van Groenigen et al., 2006), and terrestrial ecosystems often develop toward a multiply co-limited state (Harpole et al., 2011). Processes governing availability of nutrient elements, including micronutrients, were well summarized by Marschner and Marschner (1995). Few ecosystem-scale models take into account micronutrients, but P has increasingly been included in such models, particularly those addressing soil formation over multicentury or longer time scales (Taylor et al., 2011). As well as predicting the availability of individual elements, it is important to consider how interactions among nutrient availabilities can determine plant production. The concept that nutrients are used more efficiently when other nutrients are in greater supply has been implemented in models such as QUEFTS (Janssen et al., 1990). The most appropriate approach to modeling nutrient interactions may vary with the ecosystem and with data availability—a law-of-the-minimum approach (Liebig, 1840) may be adequate for agricultural systems, whereas concurrent limitation may be a more appropriate concept for more natural systems (Rastetter, 2011).

Examples of biogeochemical models at the larger scale are given in “Linking Soil Modeling Platforms with Climate, Ecology, and Hydrology.” In summary, the aspects of modeling nutrient cycling that currently offer the most scope for improvement are: interactions between litter composition and intrinsic soil properties in determining mineralization rates, links between rapid turnover of organic matter and the slower processes that determine soil development, links between nutrient availability and transport models, a focus on modeling aspects of nutrient availability that can be measured, direct effects of plants and mycorrhizae on mineralization, and interactions among nutrient elements.

### Biological Activity

Soils are home to 25% of all living species on Earth (Turbe et al., 2010) and contain a vast amount of genetic diversity, mainly derived from microbes but also plant roots (Torsvik et al., 1990; Torsvik and Ovreas, 2002). Soil biological activity derived from genetic diversity is a critical supporting ecosystem service because of the diverse metabolic pathways encoded in microbial DNA (Daniel 2004; Ferrer et al., 2009; Chan et al., 2013). These pathways include antibiotic production and resistance as well as other medically and industrially relevant natural products (Handelsman et al., 1998). In both managed and unmanaged systems, soil biological activity and genetic diversity supports emergent ecosystem services, including soil nutrient cycling, plant productivity, soil formation, and carbon storage (van der Heijden et al., 2008; Singh et al., 2010).

Despite the importance of soil biological activity, we currently lack adequate tools to predict rates of biological processes in specific soil environments or link genetic diversity to soil ecosystem functioning. Many empirical studies have begun to make this link (Hawkes et al., 2005; Prosser and Nicol, 2008; Mackelprang et al., 2011), the large number of interacting biological and physical processes pose a key challenge for modeling soil biological activity (Blagodatsky and Smith, 2012). Even at very small scales, many thousands to millions of distinct genotypes (or operational taxonomic units, OTU) may inhabit a single gram of soil (Torsvik et al., 1990; Curtis et al., 2002; Schloss and Handelsman, 2006; Or et al., 2007). Genetic diversity interacts with environmental heterogeneity in physical and chemical properties and states (Dion, 2008). Heterogeneity occurs both in time and in space, thereby driving changes in community structure and activity of soil organisms (Torsvik et al., 1996; Curtis and Sloan, 2005; Prosser et al., 2007). For example, soil hydration status and pore-space characteristics influence microbial motility, an important trait for expansion
and survival in highly patchy soil environments (Barton and Ford, 1997; Chang and Halverson, 2003; Or et al., 2007; Dechesne et al., 2010), especially in unsaturated soils with limited advective transport (Ebrahimi and Or, 2014).

Progress in resolving soil ecological questions requires quantitative models that integrate key biophysical processes with ecological interactions at appropriate spatial and temporal scales (Prosser et al., 2007). Still, such models are not yet well developed (Todd-Brown et al., 2012). Most current models of soil functioning are based on correlations between biological activity and ecosystem functions. At scales from the soil pore (Hallett et al., 2013) to the landscape (Attard et al., 2011; Eisenhauer et al., 2012) correlations between broad measurements of biodiversity or biological activity (e.g., guilds, phyla, functional groups, nutrient cycling) and soil properties (e.g., nutrients, pH, texture, pore structure) are used to parameterize soil models (Hunt and Wall, 2002; Young and Crawford, 2004; Cazelles et al., 2013). Some of these models describe the trophic relationships between organisms, including plants (Hunt and Wall, 2002). These food web models have suggested that the relationship between biodiversity and ecosystem processes is affected by land use (de Vries et al., 2013).

A new generation of models is accounting for diversity in soil organismal traits at appropriate spatial and temporal scales (Long and Or, 2009; Allison, 2012; Crawford et al., 2011). Organisms with favorable combinations of traits in a given environment will proliferate and contribute to ecological functioning or to the formation of “hot spots,” such as within soil aggregates (Ebrahimi and Or, 2015). There are several advantages to these trait-based approaches. First, they do not require information about specific organisms. Instead, genetic or other trait information can be derived from a range of sources and used to establish trait distributions for modeling. Trait values can be assigned to hypothetical organisms from these distributions at random to represent a wide range of potential ecological strategies. The environmental conditions then determine which strategies are actually viable. Second, the traits and their interrelationships can be derived from existing genomic and metagenomics data. These datasets include rich information on functional gene frequencies and correlations (Berlemont and Martiny, 2013). Finally, trait-based models can be run in different physiochemical contexts to mimic soil heterogeneity and make predictions of ecosystem services, such as the total amount of carbon storage or rates of nutrient cycling (see “Nutrient Cycling” section). Trait-based models have been applied to predict enzyme activities, decomposition rates, and N cycling in decomposing litter (Allison, 2012; Kaiser et al., 2014), as well as the warming response of C use efficiency in soils (Allison, 2014).

In soil systems, significant progress can be made by implementing organismal traits in spatially explicit, individual-based models (Wang and Or, 2014). The question of what part of genetic diversity estimates is directly linked and shaped by present ecological conditions and what fraction is shaped by population and interspecies interactions with time remains a central challenge for modern microbial ecology (Curtis and Sloan, 2005; Martiny et al., 2006; Prosser et al., 2007). Integrating these poorly understood processes into soil models presents an even greater challenge.

**Soil Degrading Processes**

**Salinization and Alkalinization**

Salinization of soil and water resources is a chronic problem in many arid regions where ET exceeds rainfall. The expansion of irrigated agriculture with marginal water sources to meet the growing demand for food is likely to increase the range of soils impacted by salinity. A confluence of conditions ranging from the projected hotter and drier climate patterns, to increasing salt loads due to use of marginal water sources, saltwater intrusion due to over exploitation of coastal aquifers; rapid withdrawal of slowly replenishing inland aquifers (e.g., Ogallala Aquifer in the USA), and mismanagement of rapidly expanding irrigation in arid regions are expected to confound this long standing problem (Assouline et al., 2015). Land degradation and loss of agricultural productivity due to salinity and sodicity hazards are among the earliest anthropogenic ecological disasters responsible for the demise of the civilizations of Mesopotamia and the Indus valley (Hill, 1992; Van Schilfgaarde, 1994; Ghassemi et al., 1995). Additionally, in some regions the buildup of calcium carbonate modifies soil hydraulic properties through the formation of low permeability carbonate enriched soil layers. Presently, about 20 to 50% of the irrigated land worldwide is salt-affected (Ghassemi et al., 1995; Flowers, 1999; Pitman and Lauchli, 2002; Tanji, 2002). Salinity damage in agriculture is estimated at US $12 billion per year, and it is expected to increase with persistent salinization of water resources (Ghassemi et al., 1995).

Crop response to the spatial and the temporal distributions of soil water content and soil salinity is complex and not fully understood, whereas it is often the combined effects of the osmotic and capillary components of the soil solution that affects plant transpiration and crop yield (Childs and Hanks, 1975; Bresler and Hoffman, 1986; Bras and Seo, 1987; Bresler, 1987; van Genuchten, 1987; Hanson et al., 2008; Russo et al., 2009; Duffner et al., 2014). Salinization has been extensively modeled based on numerical models of water and solute dynamics in agroecosystems, for example, based on the SWMS and HYDRUS codes (‘Tuli and Jury, 2004; Mguidiche et al., 2015). However, one of the most urgent modeling challenges is to improve quantitative description of the interactions between soil water salinity and plant response. Much of the know-how in the basis of salinity management (leaching, crop selection, water quality mixing) is empirically based and derived from seasonal averages, making it difficult to generalize and adapt to changing climate and future water quality and more intensive agriculture (Assouline et al., 2015).
The standard salinity management strategies often involve mixing of waters of different qualities, the selection of salt-tolerant crops, avoidance of overly sensitive soils, and compensating for high salinity water by increasing irrigation dosage above plant transpiration demand (Russo and Bakker, 1987; Shani and Dudley, 2001; Shani et al., 2007; Dudley et al., 2008; Russo et al., 2009). The traditional approach, where the leaching fraction increases with irrigation water salinity, introduces significant risks due to increasing salt loads on groundwater resources that could reduce available freshwater at the regional scale (Assouline and Shavit, 2004; Schoups et al., 2005; Shani et al., 2005). Proper assessment of such environmental risks, and the sustainability of irrigated agriculture in such systems hinges on the ability to model and predict multi-season and regional hydrological processes well beyond the single field–single season irrigation decisions of the past.

A rapidly expanding alternative source for water irrigation in arid and semiarid regions is the application of treated effluents (TE) (Hamilton et al., 2007; Qadir et al., 2007; Pedrero et al., 2010), especially in agricultural regions near urban areas (Shuval et al., 1986). Global estimates of effluent reuse indicate that about 20 million hectares of agricultural land are irrigated with TE (Jimenez and Asano, 2008). However, the increased reliance on TE for irrigation in arid regions is often practiced with little consideration of long-term impact on soil, hydrology, and ecology of the irrigated area. The primary risks associated with TE irrigation involve high concentrations of salts, especially sodium, and of organic compounds (Feigin et al., 1991; Balks et al., 1998; Hamilton et al., 2007; Pedrero et al., 2010). Recent studies have shown that long-term effects of TE irrigation resulted in a significant degradation of soil structure and hydraulic properties due to increased exchangeable sodium percentage (Leij et al., 2004; Lado et al., 2005; Assouline and Narkis, 2011, 2013; Levy et al., 2011). Evidence from other studies have shown other negative effects related to chemical aspects (Xiong et al., 2001; Wallach et al., 2005; Lado et al., 2012), as well as human health and other ecological risks associated with introduction of pathogenic microorganisms, heavy metals, and toxic organic compounds into the soil and crop (Toze, 2006; Pedrero et al., 2010; Scheerling et al., 2010; del Mar Alguacil et al., 2012). Hence, the sustainability of a coupled agro-urban hydrological cycle, where TE is used for irrigation hingens on proper management to mitigate adverse impacts of long-term TE application to avoid potential collapse of soil ecological functions.

Soil salinity management is likely to remain a challenge in the foreseeable future, especially with the growing pressure of agricultural intensification to feed the growing world population, changes in climate patterns, and increased reliance on marginal water sources. Meeting these challenges will require multidisciplinary approaches that combine modeling tools with management strategies to ensure sustainable and safe use of irrigation water resources of variable quality. We clearly need a new generation of quantitative models that integrate key biophysical processes with ecological interactions at appropriate spatial and temporal scales.

**Erosion**

Erosion can result from the action of wind, water, and tillage. In semiarid zones, wind erosion is very significant and tillage erosion redistributes considerable amounts of soil at the field scale. However, water erosion is globally the most important and will be the focus of discussion here.

The intensification of agriculture and changes in rainfall patterns with more intense rain events may increase rates of surface soil erosion. The damage is not limited to the removal of productive soil top layer (Pimental and Sparks, 2000), but also affects surface water quality downstream (stream and lake ecology, dam siltation, and enhanced pollution by agrochemicals and colloid facilitated transport). Soil erosion is strongly connected with drivers for climate change, as the mobilization of large amounts of soil organic carbon by soil transport may significantly contribute to atmospheric CO₂ emissions (WMO, 2005). In addition, drier soil conditions associated with future climate extremes may limit rates of soil carbon accumulation, thereby reducing soil aggregation and enhancing vulnerability to wind erosion. A host of soil conservation strategies for combating land degradation due to soil erosion offer additional benefits such as enhanced soil water storage (Troeh et al., 1991; Pimental and Sparks, 2000). Soil erosion leads to significant loss of agricultural land and reduction in agricultural productivity, as soil loss diminishes soil water storage capacity, impacting crop growth and enhancing flood risk. Furthermore, soil erosion plays a significant role in the biogeochemical cycles of C, N, P and Si as it redistributes significant amounts of these elements over Earth’s surface (Van Oost et al., 2007; Quinton et al., 2010). (See also “Nutrient Cycling.”) Several reviews on modeling soil erosion have been published in the past, and the reader is referred to those papers for more information on the different concepts ranging from simple models such as the Universal Soil Loss Equation (USLE), to more complex process-based models such as KINEROS (KINematic EROsion Simulation) and WEPP (Water Erosion Prediction Project) (Merritt et al., 2003; Aksoy and Kavvas, 2005).

Soil erosion by water is a complex phenomenon resulting from soil detachment by raindrop impacts and overland flow, and transport of particles by rain splash and by sheet and channel flow (Ellison, 1944, 1945). Quantitative evaluation of erosion effects at the different scales requires modeling capabilities to deal with the complexity of the processes involved. In the different modeling approaches, the driving and resisting forces are conceptually expressed by (i) flow erosivity (an indicator of the erosive potential of rainfall and runoff) and (ii) soil erodibility (a measure of the susceptibility of soil particles to detachment and transport by rainfall and runoff). Both are state variables that respond to variations in local and regional conditions, making their evaluation the real challenge of erosion modeling. The flow erosivity requires data on
the timing and amount of runoff (Assouline et al., 2007). This is required for the nontrivial issue of modeling coupled infiltration and overland flow (Furman, 2008; Chen et al., 2012; Langhans et al., 2013). Quantitative representation of the infiltration process itself requires multiscale information of soil hydraulic properties and its spatial variations, soil surface conditions, topography, soil profile initial conditions, and boundary conditions (Assouline, 2013). The amount of sediment detached or transported either by drop impact of flowing water will be determined by the soil “erodibility,” which is controlled by a range of both static and dynamic soil properties, including soil texture and soil mechanical properties (Wischmeier, 1978; Watson and Laflen, 1986; Poesen and Nearing, 1993; Bradford and Foster, 1996; Römken et al., 2002; Assouline and Ben-Hur, 2006).

Because of the multiscale nature of erosion, one can either focus on the microscale and consider soil particle detachment by rain splash and sediment transport using a process-based approach (Eckern, 1950; Rose, 1960; Lane, 1982; Diaz et al., 2008) or use an empirical macroscale approach (Pelletier, 2012). At the macroscale, the most commonly used quantitative expression of soil erosion continues to be the multiplication-of-factors type empirical equation, as proposed by (Neal, 1938), where soil loss is a function of the product of soil erodibility and rain erosivity (Wischmeier, 1978; Meyer and Harmon, 1989; Kinnell and Wood, 1992; Kinnell, 1993; Zhang et al., 1998). Following this approach, soil erodibility is considered an intrinsic soil property independent of rainfall and slope conditions (Lane et al., 1987). However, soil erodibility has been found to be dependent on infiltration and runoff (Nearing et al., 1990; Kinnell, 1993) and to change with time during the rainfall event (West, 1988; Assouline and Ben-Hur, 2006). Soil erodibility also varies over the long term due to feedbacks between erosion and soil properties (Govers et al., 2006). Another major problem with current macroscale assessments is that the procedures used for upscaling are sometimes inadequate, which may lead to a significant overestimation of erosion rates (Cerdan et al., 2010; Quinton et al., 2010).

Relatively little attention has been given to the modeling of soil transport across the landscape, in connection with soil, nutrient, and carbon delivery to stream and open waters. Whereas spatially distributed sediment routing using transport and deposition laws may offer better perspectives to understand sediment delivery, such modeling approaches have been relatively simple (Van Rompaey et al., 2001) and need further improvement to fully account for the complexity of real landscapes. Mitigating and controlling erosion require advance modeling tools to evaluate the appropriateness and efficiency of alternative approaches and methods.

**Compaction**

Soil compaction caused by human activities that reduces soil pore volume has been recognized as a worldwide problem (Bridges, 1992; Soane and van Oosterwijk, 1995). Compaction affects soil fertility by reducing water and airflow, which alters the soil’s biological activity and redox potential, induces changes in Fe mobilization and CH$_4$ emission. These changes can turn soil into a source for environmental CH$_4$ instead of a sink. Furthermore, the platy structure caused by soil compaction reduces plant rootability. Compaction also decreases water infiltration, which increases water runoff, soil erosion, and the likelihood of flooding and debris flow. Efficient protection against unwanted soil compaction requires knowledge of the mechanical processes and properties of structured, unsaturated soils. Although compaction occurs naturally during soil formation (see “Soil Formation”), the majority of soil compaction studies assess the anthropogenic impacts that cause compaction, such as tillage, vehicle and animal traffic, or forest clear-cutting with heavy harvesting equipment. All soil deformation processes affect ecosystem services and soil functions in the short term and some, such as those involving irreversible dewatering and compaction of clay, in the long term as well.

Soil compaction models use empirical (simple cause–effect relationships), semi-empirical (pedotransfer functions), and process-based approaches (Keller et al., 2013). Process-based compaction modeling is generally a three-step approach. The first step describes the load situation (e.g., pressure distribution at the soil surface under the wheel or track of a vehicle). The second step quantifies the change in the stress field within the soil due to the load applied to the soil surface. The third step uses constitutive relationships to quantify soil deformation as a result of the change in the soil stress field. These three steps are typically incorporated into an analytical (Soehne, 1953; Soehne, 1958; Horn, 2003; Van den Akker, 2004; Keller et al., 2007), or numerical model (Richards et al., 1997; Berli et al., 2003; Peth et al., 2006).

Recently, progress was made toward improving the characterization of the pressure distribution at the soil surface (Gysi et al., 2001; Keller, 2005; Lamandé et al., 2007), evaluating the different stress transfer models within the soil (Défossé et al., 2014), and determining soil constitutive relationships (Horn, 2003; Keller and Arvidsson, 2007; Berli et al., 2015). This progress allowed for improved process-based compaction modeling that used a comprehensive framework to describe stress-deformation behavior due to vehicle traffic. Although most compaction research is being done at the bulk (centimeter) scale, recent advances in nondestructive imaging (microcomputed tomography, μCT; neutron tomography; and nuclear magnetic resonance imaging, MRI) and numerical modeling with high-performance computing have allowed for compaction research at the pore scale (Berli et al., 2006; Eggers et al., 2006; Berli et al., 2008; Peth et al., 2010). Additionally, more soil information has become available because of georeferencing and global positioning systems (GPSs) that allows for soil compaction modeling at the field scale using pedotransfer functions. Horn and Fleige (2003) developed pedotransfer functions to estimate compaction sensitivity based on bulk density texture, organic matter content and soil structure as
well as moisture status. Horn and Fleige (2003) also addressed the changes in physical soil functions that were related to soil surface loads, for example, due to vehicle traffic (Duttmann et al., 2014). Assouline (2006a,b) extended models for the soil water retention and hydraulic conductivity curves to account for structural changes in soils resulting from changes in porosity, enabling the prediction of the hydraulic properties of compacted or tilled soils.

Despite the considerable progress in soil compaction modeling since Soehne’s early work (Soehne, 1953, 1958), challenges remain. For example, we have only a very limited quantitative understanding of soil structure and dynamics and how they influence the physical and mechanical processes and properties of soil (Logsdon et al., 2013). Although the description of soil stress-deformation behavior has largely improved, the impact of soil deformation on soil hydrological processes, soil chemistry, and soil biology is still not well understood. Another limitation is that classical soil mechanics were developed for mostly static loads, whereas most soil compaction is caused by dynamic loads, such as soil deformation under a rolling wheel. The differences between compaction caused by static and dynamic loads were studied only recently (Wiermann et al., 1999; Ghezzehei and Or, 2001). Finally, there is a huge gap in upscaling soil compaction properties and processes measured in the laboratory to the field scale, as well as understanding the effects of field-scale compaction on hydrological and ecological processes in the landscape. For an ecosystem-scale soil model, we suggest that a simplified semi-empirical soil compaction modeling approach would likely be the most effective to improve the quantification of soil ecosystem processes and identify the key challenges.

### Soil Modeling and Ecosystem Services

In this section we will deal with the role of soil models in understanding, quantifying, and delivering ecosystem services. We focus on two groups of ecosystem services as outlined in Fig. 1, that is, regulating and provisioning services. Regulating services include climate regulation and recycling of wastes and buffering and filtering capacity of soils; provisioning services include biomass production for food, fiber, and energy, as well as soil as habitat and physical support. We discuss the role of soil models to determine the importance of the different soil properties, as affected by the different soil processes, for the different ecosystem services. At the end of this section, we formulate five key challenges on soil modeling and ecosystem services (Table 4).

### Regulating Services

#### Climate Regulation

Climate regulation may be assessed in terms of the time scales of its regulatory function. For example, at hydrological short time scales soil water storage affects various climate patterns (e.g., rainfall events, droughts, heat waves) (IPCC, 2007), whereas for the longer term, soil serves as a sink or source of greenhouse gases (GHG) through levels of carbon sequestration (Smith et al., 2013). Soil regulatory function could also be assessed through mechanistic feedbacks related to its properties and hydro-ecological functioning, such as effects of soil on plant communities that affect climate, surface albedo, land use patterns, and more. The inextricable links between soil and climate were highlighted in the section on soil formation and have been quantified in various quantitative models for soil formation. For purposes of this review, feedbacks from soil processes that modify climate processes constitute soil’s primary regulatory role. Soil water storage features prominently in the definition of droughts (Alley 1984; Dai et al., 2004) and is considered an important factor in observed extreme heat waves (Jaeger and Seneviratne, 2011; Seneviratne et al., 2014). A recent study (Trenberth et al., 2015) has argued that the omission of soil processes (water content) in climate models seriously hampers their ability to explain the origins of a range of climate extremes ranging from droughts to floods and heatwaves.

Soil properties control soil evaporation dynamics and transition to stage 2 evaporation (Or et al., 2013), a short-term process with significant surface energy balance ramifications. On decadal to millennial time scales, the most important aspect of soil climate regulation is the soil’s role as a source or sink of carbon and other GHG (Smith et al., 2013). This is linked to the amount and stability of estimated soil carbon stocks that vary with soil properties and function (also with land use and climate). Changes in soil surface temperature affect the fate of carbon stocks in arctic regions and within a relatively short period, large tracts of land may become significant sources of GHG at high rates, for example, due to the rapid thawing of permafrost soils in northern latitudes (Schuur et al., 2015). On geologic time scales, rock weathering and formation of soils play a substantial role in supporting vegetation,

### Table 4. Key challenges to soil modeling and ecosystem services.

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Key challenge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Challenge 1</td>
<td>To include soil structural dynamics in the prediction of the soil’s buffering and filtering capacity.</td>
</tr>
<tr>
<td>Challenge 2</td>
<td>To link soil habitat and biodiversity modeling to soil function models.</td>
</tr>
<tr>
<td>Challenge 3</td>
<td>To model soil formation processes over long time scales.</td>
</tr>
<tr>
<td>Challenge 4</td>
<td>To upscale rhizosphere processes and soil–root interactions to the field scale.</td>
</tr>
<tr>
<td>Challenge 5</td>
<td>To integrate new understanding of structure dynamics, chemical and biological processes into operational biosphere models, such as those predicting GHG emissions.</td>
</tr>
</tbody>
</table>
Soil management practices, such as tillage and land clearing (forests and grasslands), are among the main human activities that have significantly increased CO₂ emissions in the past centuries, with much of the emissions mediated by soil microbial processes. Additionally, the increase in fertilizer application to boost crop production (part of what is known as the Green Revolution), has resulted in significant releases of nitrous oxides to the atmosphere, thereby reducing nutrient use efficiency and directly contributing to global warming. Vinken et al. (2014) estimated that one-fourth of soil NOₓ emissions come directly from applied fertilizers. For natural systems at the northern lower latitudes, it is expected that soil warming and melting of permafrost will result in positive feedbacks of unknown magnitudes (Schuur et al., 2015). In general, wide ranging estimates of negative feedbacks are projected with rising temperatures that could decrease net primary production. Hence, to understand the role of GHG emissions and to mitigate their adverse impacts, the soil community must endeavor to study the integrated soil system by linking physical, chemical and biological processes, and their variations with future climate patterns and introduce state-of-the-art knowledge on soil processes in existing and operational terrestrial biosphere models (Fisher et al., 2014). In particular, the assessment of the impact of management and land use practices on GHG emissions requires models that are based on a fundamental understanding of these processes. However, there are substantial deficits in presently used models both in terms of appropriate parameterization and with respect to the underlying processes (see “Nutrient Cycling” and “Biological Activity” sections above). When considering regional soil carbon balances, one must take account of changes caused by soil erosion and soil formation (longer time scales) that affecting the soil organic matter pool and the balance between its decomposition and sequestration (Lal, 2014; Amundson et al., 2015).

Soil models for climate regulation are listed in Table 2. Advanced soil modeling platforms offer a way forward that systematically uses available knowledge and considers and incorporates feedbacks (climate, soil biology, social aspects) to yield better understanding and predictive capabilities of integrated soil systems (see “Toward a Soil Modeling Platform” section). Integrated modeling approaches informed by climate scenarios and feedback provide the necessary know-how for adapting agricultural and natural ecosystems to changing temperatures and soil moisture regimes that affect plants and crop yields as well as soil ecological functioning and long term sustainability. These aspects are further discussed in the sections “Linking Soil Modeling Platforms with Climate, Ecology, and Hydrology” and “Linking Soil Modeling Platforms with Crop and Biomass Production.”

### Buffering and Filtering

We may define the buffering capacity of soil as including processes that involve storage and transformation of chemicals, including both anthropogenic and biogeogetic substances. Soil buffering is crucial with regard to the filtering capacity of soil, that is, the soil’s capacity to temporarily retain chemicals from emission to the atmosphere or groundwater. Addition and removal of chemicals disturbs the state of a soil, affecting biota as they require sufficiently stable conditions; however, such disturbances can be countered by biogeochemical processes. The modeling goal is to quantify the extent and spatiotemporal variability of such buffering.

All soil related processes are connected with soil buffering and filtering. Relevant physical processes concern the exchange of carrier fluids, such as water and gas with groundwater, surface water, and atmosphere, as well as by physical filtration at phase interfaces, whereas important biogeochemical processes are chemical ad- and desorption, precipitation—dissolution, and transformation. In addition, biological processes, like in the rhizosphere and biofilms may play an important role in filtering and buffering and have not been explicitly considered in modeling until recently (Or et al., 2007; Schimel and Schaeffer, 2012). Soil clay minerals; Fe-, Al- and Mn-hydroxides; organic matter; and carbonates play a major role in soil’s buffering and filtering capacity. Because soil organic matter is a major sorbent for many important chemicals, buffering is intensively linked with the major cycling of N, P, and C. Major inputs of, for example, nitrogen to the soil system may affect the soil’s pH, leading to acidification and changes in its buffering capacity (Guo et al., 2010; Tian and Niu, 2015). Connected with the unsaturated soil zone is the capillary fringe at the groundwater table. Since the capillary fringe is characterized by steep gradients in terms of hydraulic state variables and chemical (e.g., redox potential) and biological conditions, it involves both different processes and different rates of interactions with regard to buffering and filtering than the vadose zone. Moreover, this biogeochemically important transition zone changes very dynamically with time and depth (Winter et al., 2015). Yet, our understanding of this important zone between the vadose zone and the groundwater is still limited, requiring more intensive research and an more improved incorporation of capillary fringe processes in soil models.

Significant advances have been made during the past decades in understanding, quantifying, and modeling of buffering and filtering processes. General mineral equilibria models have been extended with validated ad- and desorption models for specific groups of solutes, such as metals (Zhang et al., 2012; Duffner et al., 2014). Interaction between soil components is crucial for quantifying buffering and filtering; inorganic and organic components might compete either for sorption sites or for forming aqueous complexes increasing solubility or decreasing sorption. A number of numerical tools have been developed during the last decade.
accounting for these interactions, mainly based on principles of thermodynamic equilibrium (Steefel et al., 2014). The generic nature of these tools allows for implementing complex conceptual models for fate and transport (Jacques et al., 2008; Leterme et al., 2014; Thaysen et al., 2014), but these models generally lack kinetics, as well as the inclusion of physical nonequilibrium conditions. This includes nonequilibrium of water–air dynamics, as these interfaces control interactions and access to sorption sites, duration of interactions and local equilibrium assumption (LEA) validity, and biological activity. Much of that dynamics is caused by soil heterogeneity, such as preferential and bypass flow. Many advances have been made in modeling soil heterogeneity both explicitly by Bellin et al. (1993); Roth (1995), as well as implicitly by Beven and Germann (2013).

Linking inorganic and organic biogeochemistry seems crucial for understanding the fate of many solutes. For example, some heavy metals form strong complexes with dissolved organic matter, as described in Fig. 3 for Hg (Leterme et al., 2014). Whereas modeling of inorganic chemical biogeochemistry often addresses specific components (e.g., heavy metals) and equilibrium relationships, models for biogeochemical N, P, K, and C typically emphasize conversion rates such as for organic matter and nitrogen mineralization. For cases where the organic matter pool may change significantly, with increasing occurrences of drought or water logging with associated redox potential changes, links between organic and inorganic interactions must be investigated. The kinetics of abiotic soil chemical changes also requires attention (Werner Stumm 1995; Schroder et al., 2008). In addition to the kinetic behavior of soil chemical processes, soil biological processes show similar rate-limited behavior, which is most likely controlled by chemical and structural soil properties. In fact, whether certain biological processes (as denitrification) occur at all, depends on the presence of the necessary microbial populations. In addition, bioavailability of contaminants for microorganisms affects the leaching behavior in essence (Beltman et al., 2008).

As soil models might be applied on long time scales for persistent contaminant, buffering and filtering cannot be independent from soil formation structural dynamics (see “Soil Formation” above) as these determine flow paths and availability of reactive sites. In summary, integrating physical aspects of nonuniform flow and solute transport with chemical and biological processes will remain a prominent focus of soil-modeling research.

Recycling of Anthropogenic Waste

Many human activities produce waste that is often released to the soil, such as chemotoxic and radioactive elements, toxic organic compounds, and potentially harmful living organisms and viruses. Waste inputs range from feedlots dung and farm animals, irrigation by wastewater nonpoint pollution by atmospheric deposition, accidental spills to deliberate dumping of industrial byproducts in highly engineered waste landfills. A specific pathway is soil amendment to reduce metal leaching or to control CO₂ sequestration (Campbell et al., 2006; Abril et al., 2008; Thaysen et al., 2014). Supporting processes such as limiting water flow through waste zones, sorption of compounds, and biological degradation help to regulate contaminant release to the biosphere by dilution, dispersion, retardation, and decay (e.g., see “Water Cycling” or “Buffering and Filtering” sections). This ecosystem service aims at quantifying the soil’s contribution to protect human health. Related examples of available models are listed in Table 2.

Impacts of soil contamination, waste disposal, or site remediation are typically assessed with risk assessment (chemotoxic compounds) or radiological impact (radioactive waste) models. Although the safety or protection provided by a disposal system is primarily focused on isolation and containment, quantification of dilution and dispersion and bioaccumulation in soils systems is highly relevant for impact calculations by biosphere models (Smith et al., 2014). Particularly within the framework of radiological impact studies, time scales could be several tens of thousands to hundreds of thousands of years.

Generally, engineered covers are put in place in typical landfills with hazardous materials. For near-surface disposal systems for low-radioactive waste disposal, as well as for high-radioactive wastes (Rosenberger, 2009), cement-based structures are buried under an engineered layered system of natural materials (Flach et al., 2007). Although covers could have an isolation function, protecting humans and other biota from the waste, their main functions are related to provide a stable physical and chemical soil environment for the waste and to limit water flow into the waste zones. Stable chemical conditions are related to durability in physical terms (e.g., cracking, increase in permeability) and chemical terms (sorption and solubility) affected by detrimental geochemical processes in the cement-based system (Glasser et al., 2008; Wang et al., 2013). Geochemical degrading and leaching processes are driven by soil pore water composition (Jacques et al., 2010) and thus different soil processes, such as weathering, microbiological,
and chemical processes (e.g., oxidation of pyrite in clay barriers), play a crucial role. The engineered barrier will also limit the water flow through the waste zone. The properties of the engineered barriers could be optimized to favor the evaporative capacity of the barrier, that is, increasing water holding capacity of the top water to promote ET or the divergence capacity by increasing lateral flow.

When contaminants are released into the soil, their transport and fate are governed by similar physical, chemical, and biological processes and pose similar modeling challenges as described for both the buffering and filtering regulating services. The main variable of interest is the flux across environmental compartments, such as the groundwater, biosphere, and atmosphere. A particular challenge is the development of a soil-like profile in the engineered barrier that alters its relevant physical, hydrological, chemical, and biological properties thereby altering their required performance. For that purpose, long-terms field experiments of years to decades (Albright et al., 2004; Nyhan, 2005) must be combined with natural or archaeological analogs (e.g., burial tombs) to benchmark conceptual and mathematical models. To deal with extremely long time scales, models should be able to incorporate long-term changes in climate, landforms, and other relevant boundary conditions. Integrated methodological approaches need to be developed to verify such models beyond the time-scale of instrumental observations, for example, by including proxy variables serving as (paleo) indicators of past hydrological conditions (e.g., vegetation, soil, or historical archives; Zwertvaegher et al., 2013; Zlinszky and Timár, 2013). As in simulating soil formation, many input variables are uncertain since they are in essence unknown for future conditions. Nevertheless, soil waste modeling as described herein requires the same kind of scenario-like quantification, as well as collaborations with related modeling communities.

**Provisioning Services**

**Biomass Production for Food, Fiber, and Energy**

By providing and storing nutrients and water, as well as serving as mechanical support for plants, soil plays a central role in biomass production. Soil also provides biochemical support for plant-essential symbionts. Optimizing crop and biofuel production relies on a thorough understanding of plant requirements, soil water and nutrient availability, and on plant uptake mechanisms. This can be partly achieved via experimental work, but modeling is needed to investigate complex interactions and feedbacks between bulk soil, rhizosphere, and plant systems under environmental constraints. Examples of models addressing this ecosystem service are listed in Table 2.

Plants change their bulk soil environment to maximize nutrient and water availability, affecting nutrient and water cycling (see “Water Cycling” above). Interacting biological, chemical, and physical processes affect crop root uptake and production (Lynch, 2007; Hinsinger et al., 2009; Richardson et al., 2009; Den Herder et al., 2010; Smith et al., 2011), especially under limiting conditions. The elements most often limiting to production are the macronutrients N, P, and K, although growth may be limited by supply of any of the essential elements. Many soil processes are directly affected by plant activities, especially in the rhizosphere. Because of soil–root–microbial interactions, the biophysical and chemical properties of the rhizosphere are different from those of the bulk soil.

To meet the crop nutrient demand, nutrients must be transported from the bulk soil into the rhizosphere toward the root surface (Marschner and Marschner, 1995). The most simple single root uptake model considers soil nutrient transport by convection and diffusion, desorption of nutrients from the soil solid phase, and uptake at the root surface, as by Michaelis–Menten kinetics (Darrah et al., 2006). For nutrients of low mobility, uptake models include root hairs, root exudation, and arbuscular mycorrhizal fungi in either rhizosphere scale models (Schnepl et al., 2008a,b; Schnepf et al., 2012) or in root system scale models (Tinker and Nye, 2000; Roose et al., 2001; Schnepf et al., 2012). In addition, nutrient uptake models have been coupled with water flow models (Somma et al., 1998; Roose and Fowler, 2004). Rhizosphere modeling includes root-induced changes in soil hydraulic properties through mucilage exudation and related effects on water and solute dynamics in the root zone as presented by Carminati and Vetterlein (2013). However, the release of rhizodeposits by roots and associated microbial activity enhances soil organic matter decomposition (Kuzyakov and Domanski, 2000) and would require the inclusion of microbial and carbon dynamics (Darrah, 1991).

Besides nutrients, plants also need water. The adequate description of water stress onset and water uptake distribution in soil is crucial for predicting plant growth transpiration flux and crop yield. Although we know water transpiration stream is driven by climatic demand and controlled by plant and soil, questions remain regarding the location and magnitude of controlling or regulating mechanisms for plant water flow (Lobet et al., 2014).

Bulk soil acts as a storage for water, and rhizosphere hydraulic properties control the availability of water to plants (Couvreur et al., 2014). Root segment–scale models (called mesoscopic) have been developed that explicitly solve axisymmetric Richards equation around a given root segment (Philip, 1957; Gardner, 1960; Raats, 2007) and allow one to estimate the soil hydraulic resistance. Yet, soil compaction induced by root growth, shrinkage–swelling of roots and soil leading to gap formation between soil and root, and specific nonequilibrium processes induced by mucilage, for instance, challenge these equations.

At the plant scale (called macroscopic scale) different approaches exist to account for root water uptake. Typically a sink term is included in the Richards equation, which includes soil resistance, plant root distribution, climatic demand, and sometimes a compensation term (Javaux et al., 2013). The challenge is to find
a mathematical expression for the root water uptake (or sink term) that best represents the key mechanisms embedded in a numerically acceptable level of complexity for the application. The upscaling of complex and dynamic rhizosphere processes can be assessed with the help of mathematical modeling (Roose and Schnepf, 2008). Different one-dimensional models have recently been proposed (de Jong van Lier et al., 2008; Jarvis, 2011). Couvreur et al. (2012) developed a simple one-dimensional solution, considering complex three-dimensional representations of root architecture. Siqueira et al. (2008) suggested solving two one-dimensional Richards equations to represent three-dimensional root water uptake. Novel root growth models and tomography techniques have recently allowed the development of three-dimensional models (Dunbabin et al., 2013) that explicitly represent the root system architecture and connect it to a sink term (see Fig. 4 with an example of a three-dimensional simulation of root water uptake using the R-SWMS model (Javaux et al., 2008).

At larger scale, the availability of many different models for root water uptake translates into high uncertainty in predicting transpiration. For land surface models, Wang and Dickinson (2012) showed that the ratio of transpiration to total ET ranged from 0.25 to 0.64 for 10 widely accepted models, with an average of 0.42. This uncertainty is due to the poor representation or validation of root water uptake modules, in particular under dry conditions (Li et al., 2013; Canal et al., 2014) and in terms of compensation mechanisms (Tang et al., 2015).

An additional modeling challenge is to link soil-root zone processes at the rhizosphere scale to the spatially variable landscape scale (Katul et al., 2012). Land surface or crop models typically have a grid size between 100 and 100 km², in which plants uptake is modeled in zero or one dimension for the sake of simplicity and computational efficiency. In zero dimension, when spatially explicit information is not available, the effects of soil properties on nutrient and water uptake is treated simply by considering total availability and access of soil water and nutrients in the soil. More advanced crop models apply one-dimensional soil modeling, using a simplified water balance and simple root depth models (Gerwitz and Page, 1974) and thereby neglecting spatial variations in soil water or nutrient content and uptake rates. However, in spatially variable soil-root condition, the one-dimensional assumption does not hold and may lead to erroneous results of ET and crop yield, especially in soil-stressed conditions (Roose and Fowler, 2004).

Soil Physical Support
Most terrestrial ecosystems rely on soil for their physical support and stability. The functional design of plant roots is optimized for sufficient anchorage to the ground (Coutts, 1983, 1986; Coutts et al., 1999). Large trees and perennial shrubs especially have root systems that are intimately linked with the soil underneath, which enables them to support the enormous weight of their own biomass and external loads (such as animals and snow) as well as dynamic stresses from wind, debris flow, and surface runoff (Stokes et al., 2014). Soils also bear the weight of all terrestrial animals and provide habitat to burrowing animals including rodents, birds, and insects. At finer scales, soils provide physical support to microbial communities. The highly modified environment in the rhizosphere as well as biological soil crusts in many desert ecosystems provide stable microstructure that serves as a habitat for microbial communities. In summary, physical support service provided by soils is an essential ingredient for the health and sustainability of terrestrial ecosystems. Soils also provide direct support services to engineered structures as well as human activities. In many places, soil—in the form of mud bricks and dirt roofs—literally serves as a physical shelter. Likewise, unpaved dirt roads and paths are vital access routes and essential in management of natural resources.

A soil is able to provide the above stated services when its strength is sufficient to support the stresses exerted on it, yet not too strong to resist necessary deformations, such as for root growth and animal burrowing. Therefore, a great deal of research related to soil as a physical support system has been directed toward understanding distribution and propagation of stress and strain in soil, as well as in quantifying the underlying rheological characteristics (including elasticity, plasticity, and viscosity) (Baumgarten et al., 2013; Hallett et al., 2013). The ability of heavy clay soils to provide physical support may also be compromised when they swell and shrink on wetting and drying, respectively. In the United States alone, expansive soils cause billions of dollars of damage to civil structures (Thomas et al., 2000). However, we need to include the aspect of rigidity in all our modeling approaches also because the shrink–swell processes alter the reference volume and may result in a complete overestimation of, for example, flux processes (Horn
et al., 2014) Similarly, drainage of peatlands and soft organic soils causes substantial consolidation that adversely impacts their physical support service (Schwarzel et al., 2002). The latter is further exacerbated by progressive organic matter loss from drained soils (Dawson et al., 2010).

To gain a full perspective on the soil’s physical support services, we need to have quantitative understanding of the following key aspects of soil strength: (i) mechanisms of support in relation to specific soil strength parameters, which includes mainly tensile strength, compressive strength, and resistance to stress dependent shearing; (ii) properties that define shrink–swell potential of clays and compressibility of peats; (iii) soil strength thresholds relevant for the physical support functions; (iv) temporal dynamics of soil strength and its relations with soil moisture and temperature; and (v) spatial variability of soil strength at multiple scales; and (vi) alteration in physical support through the growth or degradation of biological structures such as plant roots.

Fundamentally, the ability of soils to provide physical support functions is a product of interplay between stabilizing and destabilizing processes. The key stabilizing processes include: soil aggregation, which is mediated by a variety of physical, chemical, and biological processes; cementation by mineral deposits; and stabilization by burrowing animals. Soil stabilization is continually countered by destabilizing processes, including shearing forces, dynamic mechanical stresses, static loads, and slaking.

Theories and models of soil as a physical support system must involve the following key elements: (i) basic theories and modeling capability concerning mechanical strength, stress, strain, and their distribution in soils, which are generally well understood for most soil conditions over wide scale ranges; (ii) reliable techniques to quantifying stresses and strains, which are also well developed, especially for stress propagation under traffic; (iii) combined physical, chemical, and biological processes as the most influential parameters to strengthen soil systems, including the dynamic stress strength changes due to hydraulic processes also in mechanical theories; and (iv) quantitative understanding of particle-scale soil strengthening and further extrapolation from the interparticle to the meso- or macroscale.

The major open questions related to mechanistic understanding and modeling of physical support functions of soil include: (i) transient phenomena, including short-term elastic and elastoplastic responses, as well as transient coupled interactions between mechanical, hydraulic, and biogeochemical processes; (ii) stabilizing and destabilizing processes; and (iii) stress-dependent changes of soil hydraulic, thermal, and gas diffusion processes.

Soil and Biological Habitat
Life in soil follows much the same pattern as human life on the surface of the planet. For life to persist, soil microbes require sufficient accessible food resources, water, safe refuges from predators, and gaseous and hydraulic transport pathways through which they move (if motile) and be active. In terms of the soil geometry to provide the physical living habitat, key soil attributes are its porosity and its continuity and level of connectivity in space and time. Thus, soil pore and hydraulic connectivity, specific surface area, and tortuosity become key determinants of all processes that impact on soil life. The spatial distribution of porosity and nutrients determine distances between active microbes (and roots), whereas the connected porosity determines the rate at which soil gases as CO_2 and O_2 can diffuse between microbial active sites. Therefore, the soil water characteristic becomes one of the most important relationships in soil ecology (Assouline, 2013).

A traditional approach to understanding and modeling of the soil habitat, driven by the need to capture field relevant and observable metrics, has been coarse structural measures, together with some measure of “structure” through aggregate and pore size distribution and stability. The latter has driven the majority of research in this area despite the fact that the level of aggregation is the relevant metric to capture and understand for most soils (Young et al., 2001). Many conceptual models on aggregates exist (Six et al., 2004), but are rarely put into actual mechanistic models. However, understanding the relevance of soil architecture within as it varies over time is a difficult task due to the complexity of the processes at hand and the significant spatial–temporal soil dynamics. In addition, what delineates soil from all other porous media is the myriad of life and its impact on the soil’s physical architecture (Young et al., 2001). Soil formation was discussed in the above section by that name; the section “Biological Activity” addresses the incredible diversity and abundance of microbes.

The challenge in relation to modeling habitat space is its linking to the relevant functions. Biodiversity research in soils has failed generally to account for the soil habitat that controls many of the relevant processes that generate soil biodiversity, the probability of movement of microbes and higher organisms, and the probability of gene transfer and the impact of pathogens on crop plants. Therefore, the inclusion of the soil habitat in biodiversity modeling (Young and Crawford, 2014) will ensure evaluations of the importance of soil geometry on soil biodiversity, including effects of spatial isolation and population connectedness (Zhou et al., 2002).

Notwithstanding the difficult challenge of quantifying biological processes in any natural environment, modeling soil biological processes present specific challenges related to the complex and heterogeneous medium, limited observational capability into the opaque soil, and the wide range of scales where biological activity matters. The issue of scale is particularly difficult, as models are required to consider interactions taking place at the scale of microbial communities in pores (Young and Crawford, 2004; Or et al., 2007) all the way to root function affecting soil processes over large expanses of agricultural lands and forests (Dimitrov et al.,
Most soil processes and related soil ecosystem functions dealt with in this paper depend in one way or another on the architecture of soils, which determines the geometry and topology of the pore space inhabited by soil biota and through which water, gases, solutes, and particulate matter transit. The architecture of soils is acknowledged as being heterogeneous at many different scales, all the way from the distribution of soils across the landscape down to microscopic pore networks and the molecular structure of biogeochemical interfaces.

At large spatial scales (field to landscape scale), the distribution of soils is mainly determined by geology, topography, climate, and land use, whereas at smaller scales (pedon to pore scale) the continuous flow of energy promotes physical and biochemical structure formation. This produces characteristic soil architectures that typically change vertically along the main direction of flow and transport within soil profiles. Because of the nonlinearity of the different interacting processes of structure formation and decay, these changes are often distinct, leading to heterogeneous structures and vertical organization of soils (e.g., layered soil profiles).

An immediate consequence of the heterogeneous structure of the subsurface across spatial and temporal scales is that observed flow rates of water, gases, and solutes or the dynamics of state variables, such as soil moisture, temperature, and biological activity, typically depend on the scale of observation. Thus, models of soil processes (e.g., flow and transport or matter turnover) need to account for this heterogeneity, and we will discuss possible options and limitations in this section. A major challenge when one attempts to model physical, chemical, or biological processes in soils is the opacity of soil materials that hampers the quantification of their architecture.

An optimistic “fundamental approach” to represent a soil would be to describe it at the pore scale with, for instance, the Stokes equation describing the flux of water and air, the Young–Laplace equation describing the vapor–liquid interfaces, multicomponent transport equations with associated equilibrium relations at phase boundaries, a slew of equations for the multitude of chemical interactions, and yet more complicated representations of the microbial realm. For any reasonably sized soil volume, however, this is clearly neither possible due to the lack of detailed information and limited computing power, nor is it desirable because of the sheer flood of mostly redundant information. Thus, it is one of the recent challenges to develop more theoretical approaches that deliver correct representations of a given range of subscales and range of small-scale processes that reproduce the system’s behavior at a larger scale with a desired level of accuracy (Daly and Roose, 2015).

The general approach to gain a representation at some larger (macroscopic) scale is to average the pertinent processes at the corresponding smaller (microscopic) scale over an appropriate domain. Necessary prerequisites for this approach to work are (i) that the macroscopic quantities are robust with respect to changes in the averaging domain and (ii) that the microscopic quantities are in thermodynamic equilibrium at the scale of the averaging domain. Given the wide temporal spectrum of forcings, for example, through precipitation, such an averaging is restricted to rather small domains. The issue is further exacerbated by nonlinear...
processes like soil water flow, transport of reactive solutes, freeze–thaw cycles, or evaporation–condensation processes, which are all capable of generating sharp fronts and intricate patterns. The proper handling of such processes remains an open research question. Current engineering solutions typically involve the postulates that (i) the large-scale mathematical formulations are of the same form as those at small scales and (ii) so-called “effective parameterizations” can be found, which complement the large-scale formulations. An example is the consideration of nonequilibrium phenomena by decoupling state variables through an additional equation at the larger scale (Ross and Smettem, 2000). These effective parameters are typically gained from inverting physical numerical models. However, there is no evidence that the postulates are valid. It appears that proceeding to larger scales—to a field or even to a larger catchment—demands numerical simulations of the pertinent multiscale processes and quickly runs into supercomputer applications that include self-adaptive discretizations.

In the case of biological processes, such as microbial activity, subsurface heterogeneity fosters the coexistence of biochemical processes that cannot be captured or reproduced experimentally in homogenized materials. This is true for the concurrence of aerobic and anaerobic processes, as well as for the turnover of organic matter in general, which is promoted or hampered depending on the relative spatial distribution of soil biota and substrate. While its importance is well recognized, it is still unclear how to represent this heterogeneity in modeling biological activity and organic matter turnover.

### Model Concepts

Homogenization is a possible approach in the case when the various scales of heterogeneity are clearly separable, so that information from small scales can be transferred to larger scales in a meaningful way. In this case, small-scale heterogeneities can be averaged in time or space toward homogenized large-scale models that account for all the essential ingredients from the small-scale processes. Separable scales might rather be expected at small scales when moving from soil pores to aggregates and up to soil horizons. Here we can identify different levels of macroscopic homogeneity. Examples for homogenization include derivation of the Darcy flow and Richards equation (Daly and Roose, 2015), solute movement in the soil with dual porous structure (Zygalakis and Roose, 2012), uptake of nutrients by root hairs (Leitner et al., 2010; Zygalakis et al., 2011), and effects of exudation by cluster roots and resulting plant P uptake (Zygalakis and Roose, 2012).

If the scales of heterogeneity are interlaced and nested, which is typically the case at the pedon scale and beyond, modeling soil processes needs to be adapted to the spatial or temporal scale representing the relevant heterogeneity at this scale. The crucial question is to determine what is “relevant.” The dissipative nature of most processes may help to address this question. In some cases, perturbations at a given scale may smear out when the observation scale becomes much larger. This may not be true when the perturbations at the microscopic scale are associated with microbial activity. However, this assumption applies, for example, for the transport of solutes through the soil pore network that develops towards a volume-averaged Fickian regime once the transport distance is much larger than the characteristic heterogeneities within the flow field. Another example is the rapid drainage or filling of single pores that translate into a smooth curve, known as the soil water retention curve, at the larger scale. In both cases, the problem faced in reality is that heterogeneities at larger scales emerge before the limit of a well-defined macroscopic behavior is reached. A possible way to deal with this is to explicitly include the heterogeneity at the well-defined subscale, while heterogeneities at the sub-subscale and smaller scales are described by effective parameterizations and averaging (Vogel and Roth, 2003). Examples, where this concept is typically applied include (i) water dynamics in soil profiles, where effective mean hydraulic properties are used for soil horizons; (ii) water and gas exchange between the soil and the atmosphere, where the lateral distribution of soil types is considered; and (iii) solute transport in groundwater, where only the coarse structure of the conductivity field is explicitly considered, while smaller scale heterogeneities are integrated into an effective dispersivity length.

Concerning biochemical processes, the vast abundance of biodiversity in soils may allow for simplified representations at larger scales since biological communities and their biological potential and activity are controlled by the local site conditions and the metabolism of individual organisms in any specific part of the pore space is not relevant. This might be true for highly productive soils in humid regions. However, especially in water-scarce systems, the feedback between soil biota, organic matter, and water dynamics leads to complex patterns of system development (Jenerette et al., 2012) that are just starting to be explored.

### Exploring Heterogeneity

Several recent technologies and conceptual tools provide novel information on subsurface heterogeneity. Among these new methods are noninvasive three-dimensional methods, such as µCT, chemical imaging, geophysics, and remote sensing with platforms ranging from unmanned air vehicles (UAV) to satellites. These methods differ widely in their capability, resolution, accuracy, and precision (see “Modern Sources of Spatial and Temporal Data for Soil Modeling”). Their most interesting aspects are the scales of resolution and view. Some may be used in an undisturbed field situation, while others are only applicable in carefully prepared lab environments. Some capture the entire volume of interest, others just its surface. Furthermore, the quantity of interest is often not observed directly, but only indirectly via a proxy. This requires the development of appropriate transfer functions that are often just empirical relations that need data-intensive calibration procedures.
The final challenge in representing the functional structure of the subsurface irrespective of the target scale is the coherent integration of all the information on (i) the multiscale architecture (including the respective material properties); (ii) the process formulation for the chosen range of scales; (iii) the system’s coupling to the environment, which is typically represented as an external forcing but should also include the feedbacks to the atmosphere and/or groundwater; and (iv) the available data, which often need to be transferred into the chosen range of scales. In this context, top-down approaches can be highly attractive to make use of the multitude of available information, which will certainly increase in the near future, quantitatively as well as qualitatively. However, a bottom-up approach rooted in fundamental basic science observations is required to complement the top-down approach because ultimately the integration of the two, top-down and bottom-up approaches, and their synergy will enable the synthesizing of new scientific knowledge about soil systems. A joint analysis toward a consistent description of terrestrial systems may help an adequate representation emerge.

Formalisms for Considering Uncertainties Related to Model Choice

Uncertainties in soil models may arise on the conceptual level (model choice), on the parameter level (insufficient calibration data), through measurement errors, from stochasticity of system forcing, and from scaling issues. Multimodel ensemble simulations, (e.g., Neuman, 2003; Clark et al., 2008; Wohling et al., 2008; Gupta et al., 2012), such as Bayesian model averaging (BMA) are a promising approach to quantify these uncertainties; BMA reflects conceptual uncertainty through a weighted average of model-wise ensembles. Each model ensemble represents parametric within-model uncertainty, restricted to the available data though Bayesian updating (conditional simulation). The model weights are given by the so-called Bayesian model evidence (BE), which corresponds to $P(D)$ in Eq. [4]. The BE value expresses how good a model (including its uncertain parameters before conditioning) matches the available data (including their possible measurement errors), combined with a priori expert knowledge on model plausibility. Recently, Wöhling et al. (2015) demonstrated the advantages of BMA approaches for soil modeling. Unfortunately, the BMA approach is challenged by two issues. First, evaluating BE requires Monte-Carlo techniques to evaluate the fitting quality (on average over its uncertain parameters) of each model. This may become computationally prohibitive for models with long run times and with many uncertain parameters (requiring very large ensembles in the BMA context). As an alternative, the so-called information criteria (IC), such as the Akaike (AIC), Bayesian (BIC), or Kashyap (KIC) information criterion, are computationally much more feasible approximations to BE. However, a recent study by Schöning et al. (2014) demonstrates that IC often provide very inaccurate approximations to BE and thus can provide misleading results. Instead, the study reviews and benchmarks a list of alternative numerical schemes for more efficient computation of BME that pose many additional future statistical and numerical research questions.

The second challenge in BMA is constructing a set of competing models to reflect conceptual uncertainty adequately (i.e., to test different, plausible hypotheses of the soil–plant system behavior) and to ensure that a model sufficiently close to the “real” system is included. In many applications, however, building a model is time-consuming and expensive, or only a single system conceptualization is readily available. Even if a large set of plausible models exists, the entire set may, in hindsight, seem inadequate on comparison to extensive and accurate data sets.

Outside the BMA context, parameter-related uncertainty after calibration can be quantified through classical Bayesian inference and by Markov chain Monte Carlo (MCMC) simulation techniques (e.g., Vrugt et al., 2009b; Wohling and Vrugt, 2011). The use of MCMC is computationally more efficient than the brute-force Monte Carlo sampling required to operate BMA. Still, depending on the number of model parameters, the complexity of the problem and the data set size, MCMC can require up to $10^6$ or more model evaluations. If MCMC is not feasible, uncertainty quantification is still possible when assuming that all model parameters and measurement errors follow multi-Gaussian distributions (at least after transformation) and that the model equations can be linearized, and then using linear error propagation (Moore and Doherty, 2005). However, soil–plant models are typically highly nonlinear, so that linearized techniques must be treated with extreme care.

Because soil models often involve many state variables (e.g., soil moisture, matric head, transpiration, soil heat flux), the choice of data types for the above analyses plays an exceptionally large role. Different data types carry different information about the individual compartments and their respective processes (Vereecken et al., 2008). Therefore, the choice of data types has a large impact on the resulting model predictions, model performance, or model selection outcome, as shown by Wöhling et al. (2015). In such situations, multiobjective optimization (e.g., Marler and Arora, 2004; Reed et al., 2013) is a valuable tool to test how soil models fit to different data types (Santra et al., 2009; Wohling et al., 2013) used multiobjective optimization as a diagnostic tool to detect model structure errors and found large contrasts in the fitting quality to individual or combined data types. They also showed that an inadequate choice of calibration data sets may result in unrealistic parameter estimates and poor predictive performance, particularly for quantities that have not been included in model calibration. Soil monitoring in the past has been largely restricted to a limited set of standard observations (e.g., soil moisture), which may or may not be decisive to inform the parameter inference or model selection process. Therefore, the worth of different and new
Does Local-Scale Model Complexity Matter for Predictions at Larger Scales?

For local predictions, the processes and the parameters of the process model need to be described as precisely and accurately as possible. Due to soil heterogeneity, information that is available about local soil parameters or about state variables or fluxes that are used to parameterize the model is very uncertain. This uncertainty is propagated into uncertainty about predictions, which may therefore be imprecise. However, for several practical applications, not the predictions at a certain given site and time, but the distribution of a certain variable in a specific region over a certain period are required. For predictions of the percentile of the distribution in a region the set of conditions in the region needs to be represented as precisely as possible. This implies that the model should be able to represent the conditions in time and space that represent the distribution of conditions for the area and time period that is considered. The question arises therefore whether it is more important to have spatial and temporal coverage of information that is required to run a simplified and locally less precise model or whether it is better to use a more detailed and precise representation of the processes at a limited number of locations and time periods. The problem of the second approach is that the relevance of the predictions for the region and time period of interest cannot be evaluated based on the lack of spatial and temporal coverage of the model parameters and boundary conditions. The distribution, which is predicted based on a limited number of conditions or situations, may therefore lack accuracy.

An illustrative example is the process of pesticide risk assessment for pesticide registration (Leterme et al., 2007; Vanderborght et al., 2011). The general principles and questions may also be transferred to other soil processes and predictions. The pesticide fate parameters (sorption and degradation) often vary strongly with location, but their variation cannot be predicted or derived from other soil properties, so these parameters are often treated as stochastic parameters. Figure 5 illustrates the effect of uncertainty of pesticide fate parameters on the predicted cumulative distribution of leachate concentrations in a certain region. In pesticide risk assessment, the question arises whether a prediction with a detailed process model that requires detailed information

![Diagram](image-url)
about soil properties (including for instance a parameterization of preferential flow and transport) and temporal information of meteorological variables (rainfall data with high temporal resolution to capture rainfall intensities that trigger preferential flow) is to be preferred over a prediction with a much simpler model that considers only yearly rainfall amounts and uses information about soil texture and organic matter. The problem with the first approach is that an area-wide parameterization of a detailed model may not be possible due to a lack of data. For instance, detailed soil and weather data may not be available and the area-wide parameterization of preferential flow models still poses a problem, although recent advances have been made in the development of pedotransfer functions (see “Informing Soil Models Using Pedotransfer Functions” section) for these types of models (Moey et al., 2012; Tiktok et al., 2012). The second problem is that computational resources may still be limiting to carry out simulations for millions of scenarios that are required to represent the distribution of soil, vegetation (crop), and weather conditions and to consider uncertainties or spatial variability of stochastic parameters that cannot be mapped. A workaround for this problem is to use meta-models that are calibrated on a limited number of simulation runs that are performed using more detailed models (Tiktok et al., 2006; Stenumo et al., 2007). Such meta-models are simple regression models that make a direct link between available input parameters and the model output of interest. The structure of the regression model can be based on analytical solutions of the process model that are obtained for certain boundary and initial conditions. Since they are much simpler, meta-models can easily be used to make predictions for a large number of scenarios and conditions. This allows evaluation of the effect of stochastic parameters on the spatial and temporal distribution of the prediction of interest, which generally requires a large number of simulations. In general, stochastic parameters lead to wider distributions of predictions in a certain region for a certain time period (Heuvelink et al., 2010; Vanderborght et al., 2011). In addition, the error in meta-model predictions (lack of precision) could be treated similarly to the uncertainty due to stochasticity of the parameters. It is trivial that uncertainty about the parameters or stochasticity and lack of precision of the model may lead to large uncertainties in the predictions at a certain location. This prediction uncertainty can be reduced by determining the specific parameters at that location using for instance inverse modeling. However, the accuracy of this parameterized model to make predictions at other locations, where parameters are unknown, is small. Although the precision of predictions at a certain location might be low due to stochasticity of parameters and lack of model precision, the distribution of the predictions in a certain region is less affected by parameter stochasticity and model uncertainty when there is a large range of conditions and properties in the region and time period.

Numerical Approaches and Model Data Integration

Most of soil processes are strongly nonlinear and controlled by time-variable boundary conditions requiring numerical techniques to obtain solutions for states and fluxes. In this section, we discuss the most commonly used numerical approaches in modeling soil processes. Within model–data integration we use the terms forcing data and forcings for data used to drive a model, such as most common meteorological input including radiation, temperature, precipitation, air humidity, or wind velocity, among others. We discuss current approaches for model–data integration in the framework of operational research, data assimilation, and Bayesian methods.

Numerical Approaches

Advances in measurement technology, computing technology, and numerical techniques enable the development of models of ever-increasing levels of sophistication. Such models, capable of describing the inherent heterogeneity of soil environments, the temporal and spatial variability of boundary conditions, and the nonlinearity of involved processes and various constitutive relationships are usually obtained using various numerical techniques.

The numerical solution of the Richards equation (Eq. [1]) has always been highly challenging due to its dramatic nonlinearity. Early applications of numerical methods for solving variably saturated flow problems generally used classical finite differences. Integrated finite differences, finite volumes, and finite element methods became increasingly popular in the 1970s and thereafter. While finite difference methods today are used in a majority of one-dimensional models, finite volume methods and/or finite element methods coupled with mass lumping of the mass balance term are usually used in two- and three-dimensional models. Finite element and finite volume methods used with unstructured triangular and tetrahedral elements allow for a more precise description of complex transport domains compared to finite differences. Most popularly used vadose zone flow models (e.g., van Dam et al., 1997; Šimůnek et al., 2008) presently utilize the mixed formulation of the Richards equation and the numerical scheme of Celia et al. (1990), which possesses mass-conserving properties for both finite element and finite difference spatial approximations. Other mass-conserving numerical approaches are also available (e.g., Rathfelder and Abriola, 1994). To overcome problems of numerical stability and convergence of the numerical solution, especially for problems involving infiltration into initially dry soils, various primary variable-switching techniques have been proposed (Forsyth et al., 1995; Diersch and Perrochet, 1999; Krabbenhof, 2007). Advances in numerical techniques allowing coarser spatial and temporal discretizations are urgently needed (Vogel and Ippisch, 2008).

The numerical solution of the convection–dispersion equation (Eq. [2]) presents a different challenge, due to its simultaneous parabolic
Advances in computing technology allow development of codes for solving the convection–dispersion solute transport equation, which can be broadly classified into three groups: (i) Eulerian, (ii) Lagrangian (or method of characteristics), and (iii) mixed Lagrangian–Eulerian methods. In the Eulerian approach, well suited for parabolic equations, the transport equation is discretized by means of a usual finite difference or finite element method using a fixed grid system. For the Lagrangian approach, well suited for parabolic equations, the mesh moves along with the flow, or remains fixed in a deforming coordinate system. A two-step procedure is followed for a mixed Lagrangian–Eulerian approach. First, convective transport is considered using a Lagrangian approach in which Lagrangian concentrations are estimated from particle trajectories. Subsequently, all other processes including sinks and sources are modeled with an Eulerian approach using any finite element or finite difference method, leading to the final concentrations.

For certain problems, such as convection-dominated transport or the transport of steep fronts, the Eulerian method can lead to artificial oscillations (under or over shading) or numerical dispersion due to truncation errors of the discretization (Neumann et al., 2011). Although these numerical oscillations can be minimized by the use of upstream weighting, this can lead to considerable numerical dispersion. Since in many applications the presence of even minimal oscillations (such as negative concentrations in reactive transport models) can corrupt the solution, there exists a large family of schemes that aim to suppress such oscillations. These schemes, which use various types of flux/slope limiters, are commonly referred to as total variation diminution schemes (e.g., Leonard, 1991), and they dramatically improve the solution near steep gradients and remove under- and overshoot problems by preserving local monotonicity.

A system of linear equations, resulting from discretization of governing partial differential equations, is usually solved using different types of iterative matrix solvers, such as the preconditioned conjugate gradient method (e.g., Herbst et al., 2008a), the generalized conjugate residual method Orthomin (e.g., Mendoza et al., 1991), or algebraic multigrid methods like SAMG (Jones and Woodward, 2001; Stuben, 2001).

Advances in computing technology allow development of codes that significantly decrease the computational time by distributing complex large-scale problems over multiple processors working in parallel (e.g., Vereecken et al., 1996; Hardelauf et al., 2007). Standard parallelization approaches, such as MPI (message passing interface; Balay et al. (2015) and OpenMP (open multiprocessing), are currently being used to develop codes for both distributed and shared memory platforms, (e.g., Steefel et al., 2015). Parallelization is especially valuable for reactive transport models, in which evaluation of various biogeochemical processes consumes substantially more computational time than evaluation of flow and transport processes. The principal benefit of parallelization is that highly complex simulations can be performed in hours on a massively parallel computer instead of weeks on a desktop computer. While such models are readily available for computer systems running Linux or Unix operating systems, they are not yet readily available for personal computers with the Windows operating systems.

Since most of the soil models are based on systems of partial differential equations, generic partial differential equation solvers that were originally developed in computational fluid dynamics are becoming more widely used in soil modeling. These tools offer the advantage that the model development can be separated from its numerical solution, at the same time providing highly efficient numerical solvers for different classes of problems. Examples are OpenFOAM (www.openfoam.org), Dune (www.dune-project.org/), and FlexPDE (www.pdesolutions.com) (accessed 27 Apr. 2016).

### Novel Optimization Methods and Their Application to Soil Modeling

Model predictions for flow and transport processes in the unsaturated zone are affected by systematic and random errors. This concerns model input parameters like saturated hydraulic conductance, model forcings like precipitation, model initial conditions like soil moisture content or carbon pools, and boundary conditions like the functioning of drainagee. The model itself is also affected by errors because some processes might be misrepresented and other relevant processes not included in the model (e.g., preferential flow). In addition, model parameters are not necessarily experimentally viable to measure, or perhaps data need to be transformed before it can be used within a model (inverse modeling). Temporal and spatial soil data can be expensive to collect and knowing how much data are useful for models can be hard to gauge, as discussed in “Does Local-Scale Model Complexity Matter for Predictions at Larger Scales?” Upper and lower bounds can be derived for parameters and models are used to find the best estimate for those parameter values (fitting a given data set) via the use of operational research. Operational research is a discipline that uses advanced analytical methods to help find a better solution for a problem (lower cost) or predict what may happen to a commodity/resource in the future (forecasting). The advanced analytical methods are generally in the form of algorithms which are used to find the optimal solution of a problem. The main properties of an algorithm include the run time, convergence, and function calls. These properties are different between algorithms, with each algorithm having its own strengths and weaknesses for certain types of problem. For a nontrivial problem, picking the “best” algorithm increases the chance of finding an optimal solution given desired constraints.

An optimization problem is generally of the form

\[
\begin{align*}
\text{minimize} & \quad \Phi(X_1, X_2,\ldots, X_n) \\
\text{subject to} & \quad g_i(X) \leq 0, \quad i = 1,\ldots,n_{ic} \\
& \quad h_i(X) = 0, \quad i = 1,\ldots,n_{ec}
\end{align*}
\]
for an objective function \( \Phi(X) \), with \( n \) parameters \( (X = [X_1, X_2,..., X_n]) \), \( n_{ie} \) inequality constraints \( g(X) \), and \( n_{ec} \) equality constraints \( b(X) \). The type of variables used can either be integer, continuous, or mixed depending on the problem (Winston and Goldberg, 2004).

The mathematical models used to describe water or nutrient flow or solute transport in soil and uptake into plant root systems can produce complex parameter search spaces where numerical simulations often provide the best solution. When trying to validate such models to experimental data, a set of parameters are often constrained to vary within specified upper and lower bound, ensuring a realistic solution. This often leads to a nonlinear unconstrained optimization problem that can be solved using a given algorithm.

Nonlinear unconstrained optimization methods can be split into two categories, local and global optimization methods. Local optimization methods, or decent methods, can be categorized further into zero-, first-, or second-order methods. Zero-order methods do not use any derivatives of the objective function throughout the optimization process, for example, simplex search (Nelder and Mead, 1965), Hooke and Jeeves method (Al-Sultan and Al-Fawzan, 1997), and a conjugate direction method (Powell, 1964). First-order methods take first-order derivatives of the objective function throughout the optimization process, for example, gradient descent (Guely and Siarry, 1993), quasi-Newton’s method (Dennis and More, 1977), and a conjugate gradient method (Gilbert and Nocedal, 1992). As it follows, second-order methods use second-order derivatives throughout the optimization process, for example, Newton’s method (Battiti, 1992), a trust-region method (Byrd et al., 1987), and Levenberg–Marquardt method (More, 2006). First-order derivatives give an indication of which direction to search in, whereas second-order derivatives give an indication of how far to search in a possible optimal direction. Local optimization methods, however, converge to local optima and do not necessarily perform well on the global scale, relying heavily on good initial starting points. For complex search spaces, where there are many local optimal points, local search algorithms tend to perform worse than global search algorithms due to converging early or being stuck at one of the many local optimal points.

Global optimization methods can be split into two types, deterministic and stochastic. Deterministic methods involve no element of randomness. Therefore, any change to the optimal solution comes from different initial starting points or parameters set at the beginning of the optimization process. Deterministic global optimization algorithms include Lipschitz optimization ideas (Shubert, 1972), covering methods that iteratively tighten bounds on the global solution (Hansen et al., 1991), and generalized descent methods where local optima are penalized to encourage global search (Cetin et al., 1993). Stochastic global algorithms include clustering methods (Torn, 1977), random search methods such as simulated annealing (Aarts and Korst, 1989) and genetic algorithms (Horst et al., 2002), and methods based on stochastic models such as Bayesian methods (Mockus, 1989) and Kriging (Krige, 1952; Forrester et al., 2008), which, in addition, approximates the objective function.

There are many algorithms available for use in global optimization, and models can range from having cheap to expensive objective functions, where the number of function calls from an algorithm can become an issue. Expensive objective functions in combination with a large number of function calls make certain algorithms unusable. A major concern with global optimization is the number of variables used within a model, where the greater the number, the bigger the search space and less likely a good solution will be found within a reasonable computational time. For problems with a large number of variables, approximations models can be used that sacrifice accuracy for speed. Such approximations can take the form of simple regression models (a type of metamodel) and due to their simplistic nature, drastically decrease the run time of an algorithm.

**Data Assimilation**

Traditionally, model–data mismatch is handled in the soil modeling community by inverse modeling techniques. Inverse modeling techniques adapt, for example, the uncertain soil hydraulic parameters so that observed and simulated time series of model states coincide more closely. These inverse modeling techniques are typically based on the minimization of a two-part objective function, which includes the weighted sum of squared deviations between simulated and measured states and the weighted sum of squared deviations between posterior and prior parameter values. This objective function can be derived from Bayes theorem assuming normal distributions for states, parameters, and observations. In the last decade the focus has shifted toward calculating not just one, but multiple equally likely solutions for the inverse modeling problem. The MCMC technique is a popular approach in this context (Vrugt et al., 2003). It is a flexible approach that does not require that states and/or parameters are normal distributed. However, a disadvantage is that a large number of model evaluations is needed for the characterization of the posterior probability density function (pdf), especially if many uncertain parameters are considered and in case many measurement data are available. Therefore, MCMC is often applied for the estimation of few parameters only, for example the soil hydraulic parameters of a limited number of soil horizons. The MCMC methods have become faster with multimethod adaptive evolutionary search approaches (Vrugt and Robinson, 2007; Vrugt et al., 2009a). Recent developments include multiple try sampling, snooker updates, and sampling from an archive of past states (Laloy and Vrugt, 2012). It allows the estimation of hundreds of parameters with MCMC.

An interesting alternative that has emerged in the context of soil model–data fusion is sequential data assimilation (SDA). In this case, measurement data are not assimilated in a batch approach, but sequentially, stepping through time. Sequential data assimilation is based on the Markovian assumption, which would imply...
that the sequential incorporation of measurement data instead of the batch approach does not significantly reduce the information content of the data. A further simplifying assumption which can be made in SDA for the updating step, is the normal distribution of states, parameters and data. The Markovian and normal assumptions give rise to the Ensemble Kalman Filter (EnKF) (Evensen, 1994; Burgers et al., 1998). EnKF needs much less CPU-time than the McMC approach, although the full posterior pdf also is derived. The sequential nature of the approach is especially suited for real-time predictions of, for example, soil moisture evolution. In addition, the framework is flexible for handling multiple sources of uncertainty. A further advantage is that time-dependent parameters can be estimated. The particle filter is another SDA method and does not rely on the Gaussian assumption (Arulampalam et al., 2002). However, the approximation of the posterior pdf with the particle filter requires a large number of model evaluations and is not as efficient as the EnKF (van Leeuwen, 2009).

Sequential data assimilation has been the method of choice for model-data fusion in land surface modeling for more than a decade (e.g., Reichle et al., 2002) and more recently also for groundwater modeling (Chen and Zhang, 2006). In land surface modeling, this involves updating of soil moisture contents with remote sensing information, (e.g., Dunne et al., 2007), or in situ measurements (e.g., De Lannoy et al., 2007), and updating of soil carbon pools in biogeochemistry models, (e.g., Zhou et al., 2013). Soil parameters are in general not updated in those applications. In the following, we focus on parameter estimation with SDA for soil hydrological models, which is a less studied subject. Early applications of SDA in soil hydrology are the one-dimensional synthetic experiments with the assimilation of soil moisture data by Montzka et al. (2011) with the particle filter and Wu and Margulis (2011) with EnKF. They updated both states and soil hydraulic parameters of the van Genuchten model. Montzka et al. (2013) estimated also time-dependent variables of a radiative transfer model with the particle filter and applied the filter on a site in Colorado, USA. Wu and Margulis (2013) extended their framework for the assimilation of electrical conductivity data and applied the filter to data at site in California, USA. Although these works showed promising results, other one-dimensional studies pointed to the limitations of EnKF. Erdal et al. (2014) pointed out that a wrong conceptual model of the vertical distribution of soil horizons affects soil hydraulic parameter estimation, and they suggested the inclusion of an additional bias term to improve the filter performance. Erdal et al. (2015) stressed that especially under dry conditions the pdf of pressure is highly skewed and EnKF unstable. They showed that a normal score transformation (Zhou et al., 2011) strongly improved filter performance. Song et al. (2014) estimated two-dimensional spatially distributed saturated hydraulic conductivities of the unsaturated zone with an iterative variant of EnKF. However, their work made various simplifications, like perfect knowledge of the other soil hydraulic parameters and a constant rainfall rate. Integrated hydrological models also model flow in the unsaturated zone with the three-dimensional Richards equation. First efforts are being made to estimate model parameters of integrated models with SDA. Shi et al. (2014) estimated several soil hydraulic parameters of such an integrated model, assuming a spatial homogeneous distribution. They used multivariate data assimilation with EnKF. Pasetto et al. (2015) estimated three-dimensional spatially distributed saturated hydraulic conductivities for the unsaturated zone using the integrated hydrological model CATHY (Paniconi and Wood, 1993), assuming perfect knowledge on the other soil hydraulic parameters. Kurtz et al. (2015) developed a data assimilation framework in combination with the integrated terrestrial system model TSMP (Shrestha et al., 2014) and showed in a synthetic test the feasibility to estimate three-dimensional spatially distributed saturated hydraulic conductivities of the unsaturated zone at a very high spatial resolution (both $2 \times 10^7$ unknown parameters and states). Other data assimilation studies with integrated hydrological models excluded parameter updating in the unsaturated zone because of instabilities (Rasmussen et al., 2015).

In summary, SDA is of particular interest in soil modeling for real-time applications with the need of forecasting, for example, for real-time optimization of irrigation scheduling. Such applications require often only state updating. A second important area of application in soil modeling is high-resolution characterization of two- and three-dimensional distributed fields of soil hydraulic parameters. However, we are still facing many challenges. The main obstacle is the joint estimation of distributed fields of saturated hydraulic conductivity, van Genuchten parameters $a$, $n$, and porosity. Even if enough conditioning information would be available, this is highly challenging given the strong nonlinearity and non-Gaussianity of the problem. A further problem for real-world applications is the lack of precise data. In addition, processes like preferential flow might influence soil moisture redistribution and are difficult to capture with the standard three-dimensional Richards equation. We expect an increased use of SDA in the context of soil modeling and the use of variants of EnKF that work better for the described conditions. It is clear that a successful application requires some simplifications of the estimation problem, but those should be less stringent than in many current applications. Finally, SDA is also of interest for estimating time-dependent soil and vegetation properties and provides information helpful for improving monitoring designs.

**Bayesian Approach for Model–Data Integration**

The usefulness and applicability of soil models for system characterization and science-based decision making depends in large part on the parameterization that is used to characterize the soil domain of interest. This includes (among others) the functional form and assumed spatial variability of (i) the soil water retention and hydraulic conductivity curves; (ii) root distribution and uptake; (iii) biomass, nutrients, and biological activity; and (iv)
preferential flow, as well as the assumed soil layering, and applied lower and upper boundary conditions. In principle, in situ observation and experiments in the laboratory could help determine an appropriate parameterization of the soil hydraulic parameters, presence of flow paths and layering, biologic activity, nutrient type, amount, and distribution and root characteristics. Yet, such data often pertain to a relatively small soil volume, and the parameters derived from this analysis cannot readily be used in soil models that simulate water, ecological, biological, and biogeochemical processes at much larger spatial scales. Because of the high non-linearity of the soil hydraulic functions, their application across spatial scales is inherently problematic. Specifically, the averaging of processes determined from discrete small-scale samples may not be representative of the key processes of the larger spatial domain. In addition, the dominant hydrologic flow processes may vary between spatial scales, so that potentially different models need to be used to describe water flow at the soil pedon, field, or watershed scale, as outlined in the “Heterogeneity: Aggregate to Landscape, Microbe to Forest, Grains to Ecology” section.

In recent years, Bayesian inference has found widespread application and use in the modeling of soil processes to reconcile system models with data, including prediction in space (interpolation), prediction in time (forecasting), assimilation of observations and deterministic/stochastic model output, and inference of the model parameters. Bayes theorem states that the posterior probability, \( P(H|D) \), of some hypothesis, \( H \), is proportional to the product of the prior probability, \( P(H) \), of this hypothesis and the likelihood, \( L(H|D) \), of the same hypothesis given the observations, \( D \), or

\[
P(H|D) = \frac{P(H)L(H|D)}{P(D)}
\]

where the evidence, \( P(D) \), acts as a normalization constant of the posterior distribution, so that the posterior distribution integrates to unity. The evidence (also called marginal likelihood) can be ignored during inference of the parameters, but is of crucial importance in model selection. The hypothesis, \( H \) often constitutes some numerical model, \( F(x) \), which summarizes, in algebraic and differential equations, state variables and fluxes, all our knowledge of the system of interest, and the unknown parameter values, \( x \) are generally subject to inference using the data \( D \). Latent variables can be used to specify explicitly errors in model inputs (boundary conditions). For complex soil models the posterior distribution, \( P(H|D) \) is often highly dimensional and analytically intractable, and Monte Carlo simulation methods are required to approximate the target (Vrugt et al., 2008, 2009a,b; Laloy and Vrugt, 2012).

The Bayesian approach provides a quantitative framework to treat explicitly all sources of uncertainty, including model input (boundary conditions), model parameter, calibration data, and model structural (epistemic) errors. This latter error summarizes the effects of (among others) incomplete knowledge of soil processes and system heterogeneities. Practical experience suggests that model input and model structural errors are most difficult to describe accurately. These two sources of error do not necessarily have any inherent probabilistic properties that can be easily exploited in the construction of a likelihood (objective) function. While we can assume an error model (stochastic or deterministic) for the model input (forcing data) errors, this will be purely for the sake of mathematical convenience (Gupta et al., 1998). Consequently, it is very difficult to decompose the residual error between model simulations (predictions) and data into its constituent sources, particularly in cases common to complex systems where the model is nonlinear and different sources of error interact non-linearly to produce the measured deviation (Vrugt et al., 2005; Beven, 2006). One key challenge is therefore to improve our understanding of measurement data errors at different temporal and spatial scales. This would improve considerably Bayesian inference of soil models as prerequisite for advancing process understanding. Another key challenge is to improve model calibration and evaluation methods so that they are powerful enough to diagnose, detect, and resolve model structural errors. This is key to improving our process knowledge, and thus a prerequisite for scientific discovery and learning.

In recent years, much progress has been made in the development of process-based model evaluation methods that much better extract information from the available data (Gupta et al., 2008; Vrugt and Sadegh, 2013). These methods have been developed in the surface hydrologic literature and recognize that the very construction of the likelihood (objective) function, as a summary variable of the (usually averaged) properties of the error residuals, dilutes and mixes the available information into an index having little remaining correspondence to specific behaviors of the system. This inspired Vrugt and Sadegh (2013) to advocate a likelihood-free diagnostics approach to model–data synthesis. This approach, also referred to as approximate Bayesian computation, uses summary metrics of the original data, rather than the data, \( D \) itself. By designing each metric to be sensitive only to one component of the model, any mismatch between the simulated and observed summary metrics can be directly linked to a particular process in the model. A step back to simpler boundary conditions and system heterogeneities that allow an analytical solution or analysis of the model may be a strategy to derive these summary metrics so that a large step forward can be taken when analyzing numerical model output with appropriate metrics. An alternative strategy could be to analyze the model outputs using coherence spectra, wavelet analyses, and other decomposition methods.

Thus, as community we face a large number of challenges. First, there is the challenge of improving the description of measurement data error and uncertainty at different spatial and temporal scales. This would help us to much better constrain the model input data, and consequently help us understand whether a model is fit for a purpose (input uncertainty explains model deviations from data)
models can be informed by remote sensing in different ways. For example, these can include providing information about model forcings, model parameters, state variables, and fluxes, as well as by indirect methods using the plants as “sensors” of root zone properties (Wilson, 2009). In the following, we discuss these main measurement applications separately, knowing that their role of informing a soil model can change depending on the model characteristics.

Model Forcings
Models can benefit from remotely sensed model-driving forces when in situ measurements are not available or do not capture the spatial heterogeneity. Typically, soil models are driven by meteorological measurements, which are operationally recorded by remote sensing (Sheffield et al., 2006), such as for weather-forecast applications. One example is precipitation, measured by microwave sensors on tower-based and spaceborne platforms. The Global Precipitation Mission (GPM) is an international network of satellites that provide global observations of rain and snow, building on its core satellite, Tropical Rainfall Measuring Mission (TRMM) (Huffman et al., 2007). Similarly, networks of local weather-radar systems are combined to generate area-wide precipitation maps in high spatial and temporal resolution (Krajewski et al., 2010). Another system measures land-surface temperature, retrieved operationally by TIR sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS) or the Spinning Enhanced Visible and Infrared Imager (SEVIRI) via a generalized split-window technique (Tomlinson et al., 2011).

Model Parameters
Digital elevation models (DEMs) are among the first remotely sensed data sources to predict soil characteristics. By simple landform attributes, such as elevation, slope, and aspect, in combination with geostatistical techniques, more information about a catena such as topsoil gravel content, soil depth (Odeha et al., 1994), clay content (Greve et al., 2012), erosion (Lee and Liu, 2001; Vrieling, 2006), or even soil pH (Castrignano et al., 2011) can be predicted. However, the acquisition of DEMs, typically by Light Detection and Ranging (LIDAR, Liu, 2008), Synthetic Aperture Radar (SAR, e.g., Gruber et al., 2012), or stereoscopic optical imagery (Fujisada et al., 2005) is not straightforward because raw data can contain return signals from human-made objects or vegetation rather than bare-earth targets. Nonetheless, a large variety of high-accuracy DEMs are available from local to global scale. Ground-based and near-ground based (e.g., UAV-mounted) LIDAR and structure from motion techniques are providing proximal tools for high resolution mapping of microtopography and vegetation.

Passive optical sensors operating from VIS to NIR bands typically are designed as multichannel detectors either with a few broad bands (multispectral) or with more than one hundred narrow bands (hyperspectral). A hyperspectral imaging system,
also known as an imaging spectrometer, is better able to repre-
sent the spectral response of a target soil surface and can provide
valuable information about soil properties; already a few examples
for operational application in agricultural management such as
precision agriculture exist (Ge et al., 2011). Specific absorption
features—\(\sim 550\) nm for iron oxide (Viscarra-Rossel and Behrens,
2010) \(\sim 1730\) nm for organic carbon (Ben-Dor et al., 1997), and
\(\sim 2260\) nm for clay (Lagacherie et al., 2012)—correlate well with
in situ measurements of soil properties (Ben-Dor et al., 2009; Bayer
et al., 2012; Babaeian et al., 2015).

In some soil models, few of these observed properties can be used
directly as parameters. Implementation in pedotransfer functions
(PTFs) is an alternative approach to informing soil models by these
remote sensing—derived soil characteristics (see “Informing Soil
Models Using Pedotransfer Functions” section).

State Variables

Microwave sensors, such as radars (active) or radiometers (passive),
are able to detect variables valid for upper soil layers, including
moisture (Njoku and Entekhabi, 1996; Kornelsen and Coulibaly,
2013), roughness (Davidson et al., 1998; Panciera et al., 2009), and
salinity (Komarov et al., 2002). The challenge is to disentangle the
impacts of these variables on the MW signal, to retrieve the vari-
ables separately. Typically, salinity can be neglected for most soils,
but differentiating moisture from the altering roughness effects
is a remaining challenge (Shi et al., 1997; Verhoest et al., 2008).

One interesting approach to detect variables is the combination of
measurements obtained at different incidence angles (Srivastava et
al., 2003) or different frequencies, that is, with different sensitiv-
ity to soil moisture and soil surface roughness. Use of time-lapse
MW observations and coupled-inversion or data-assimilation tech-
niques with hydrological soil models (see also “Numerical
Approaches and Model Data Integration”) also proved to be one
of the most potent venues for soil hydraulic property estimation
from local to regional scales (Mohanty, 2013; Dimitrov et al., 2014;
Jonard et al., 2015). Other approaches make use of the spatiotem-
poral variability of surface soil moisture to indirectly estimate
hydraulic properties (van Genuchten, 1980), not only for the top-
soil, but also for the root or vadose zone (Montzka et al., 2011;
Kumar et al., 2012).

Fluxes

Energy-balance and mass-conservation rules should be considered
when informing soil models by remotely sensed flux measure-
ments in the soil–plant–atmosphere continuum. Energy-balance
components, such as latent and sensible heat, or water-balance
components, such as actual ET, can be retrieved based on sur-
face-characteristic parameters (e.g., leaf area index, land surface
temperature, surface albedo) obtained by a combination of VIS
to TIR data (see also “Informing Soil Models Using Remote
Sensing”) (Bastiaanssen et al., 1998; Mu et al., 2011).

Vegetation Canopy Properties Providing
Information About Soil Status

Spatial heterogeneity of subsurface properties, such as soil moisture,
soil texture, and soil structure, as well as biochemical properties
(e.g., organic carbon, nutrient status, pH) in combination with cli-
matic conditions, are known to affect plant health (De Benedetto
et al., 2013). Inversely, indirect methods using the plants as “sen-
sors” of root-zone properties (Wilson 2009) can therefore be used
to inform soil models. Rudolph et al. (2015) presented the link
between crop-status patterns in large-scale multispectral satellite
imagery with multi-receiver electromagnetic induction hydro-
geophysical data. Moreover, Vereecken et al. (2012) analyzed the
potential of MW remote sensing to identify water stress–related
phenomena in vegetation canopies, which can be related to sub-
surface properties.

In general, several sensors and methods still make use of ground-
based manual measurements using remotely sensed parameter
maps for regionalization and pattern recognition (e.g., Lagacherie
et al., 2012), but the number of solely air- and spaceborne applica-
tions for spatial and temporal soil property estimation is limited.
Instead of regression analyses to upscale from point to regional
scale, physical models describing radiative transfer processes need
to be developed. Future technical improvements and new sensor
developments will foster this field of research.

Proximal Soil Sensing, Geographical
Databases of Soil Properties for Soil
Process Modeling

Proximal Soil Sensing

Modeling soil processes at field, catchment, and larger extents
requires access to high resolution and spatially distributed infor-
mation on soil properties. Proximal soil sensing (PSS) has the
potential to benefit soil process modeling by increasing the cost
effectiveness and rapidity of soil characterization and monitoring,
Proximal soil sensing is the acquisition of information about the object or feature of interest using equipment either in direct physical contact with the in situ object or very close to it. "Very close" means within a few meters, usually closer. In relation to soil, proximal sensing is both a very old and a relatively new discipline—old in that the earliest soil scientists relied almost entirely on visual observations of soils in the field, and new in that recent technologies have greatly expanded and improved our ability to acquire information from the soil. Application of PSS will lead to easier process-model conceptualization, parametrization, initialization, and evaluation and will reduce the time and effort required in the "transaction costs" that surround soil modeling. Examples of PSS technologies and their applications. Recent developments in remote sensing of soils include portable X-ray fluorescence (Zhu et al., 2011), apparent electrical conductivity measurements using electrommetrical resistivity tomography, (Samouelian et al., 2005; Koestel et al., 2008), electromagnetic induction (Weller et al., 2007; Saey et al., 2009; Rudolph et al., 2015), spectral-induced polarization (Slater et al., 2006), ground-penetrating radar (Huisman et al., 2003; Lambot et al., 2010), and gamma-ray spectroscopy (Rosell and McBratney, 1998; Rawlins et al., 2007) field NIR spectroscopy (Rosell and McBratney, 1998; Rodionov et al., 2014), and ion-sensitive field-effect transistors (Lobsey et al., 2010).

Adamchuk and Rossel (2010) and Rossel et al. (2011) reviewed PSS technologies and their applications. Recent developments in sensor fusion examine the possibility of linking multiple sensors with common calibration and data-analysis approaches (Kweon, 2012; Mahmood et al., 2012), which would allow researchers to capture all of the data required to set up or validate a soil process model with one set of readings. A wide, constantly expanding range of soil parameters can be estimated using PSS, including particle-size fractions (Buchanan et al., 2012), soil moisture, root density, and available water-holding capacity (Hedley et al., 2010), clay content (Waizer et al., 2007), organic carbon (Viscarra Rossel and McBratney, 2003; Stevens et al., 2013), organic carbon fractions like black carbon and particulate organic matter (Bornemann et al., 2008; Bornemann et al., 2010), and nutrients (Wu et al., 2014). In addition to measurement of parameters, the evaluation of soil processes may also be amenable to PSS techniques (Dematte and Terra, 2014). Soil parameters estimated from proximal soil sensors can be an input to a soil inference system, where properties related to transfer of water, heat, gas, or solute can be estimated (McBratney et al., 2006). This procedure would have obvious benefits for soil process modeling because it would directly capture detailed information about what is being modeled.

The integration of PSS within soil mapping, monitoring, and modeling (SM3) is an active field closely linked to the European Soil Thematic Strategy; notable examples of efforts in this area are DIGISOIL (Grandjean et al., 2010) and iSOIL (Werbani et al., 2010). Several challenges exist, including removal or accounting for the effects of moisture and soil structure from sensor readings obtained in the field. (Minasny et al., 2011), for example, provide a solution for soil moisture. Rodionov et al. (2014) expanded the solution to handling moisture and soil surface roughness for the sensing of soil organic C. The use of spectral libraries derived from dried ground samples to calibrate models that then use field-based spectra is making good progress (Ge et al., 2014). Sampling and calibration is another growth area for PSS; these are often considered separately, when in fact they are closely related. The sampling strategy used in the field or laboratory strongly impacts data availability for calibration purposes, and the calibration method employed often places specific requirements on the quantity, variability, and type of data to be used. The interaction of sampling and calibration has been studied in the iSOIL project (Nüscher et al., 2010) and in other research (Dematte et al., 2006; Brown, 2007; Sankey et al., 2008).

Proximal soil sensing techniques often produce big data that can require complex and customized analysis, whereas the priority in terms of process modeling will be to increase data availability and eliminate much of the effort required in interpreting the sensor data. Portable infrared instruments capture ultraspectral (i.e., data across thousands of wavelengths), reducing the number of data without losing useful information makes for more accessible analysis (Viscarra-Rossel and Behrens, 2010) or spectral response-based PSS. In addition, methods of applying three-channel RGB data will open up the possibility of using digital cameras and mobile phones for PSS (Viscarra-Rossel et al., 2009; Astkenhead et al., 2014). Measurement of soil-horizon characteristics, including depth of impermeable layers, is also possible with digital imagery (Islam et al., 2014). Based on hyperspectral camera records it has also been possible to provide maps of elemental concentrations for C, N, Al, Fe, and Mn for each mineral soil horizon. VIS-NIR spectroscopy also allows differentiation of organic surface layers and the assessment of their qualitative OM properties with a high spatial resolution (Steffens and Buddenbaum, 2013; Steffens et al., 2014). Digital soil morphometrics (Hartemink and Minasny, 2014) is a subfield of PSS in which the spatial variation of sensor reading within the profile is used to enhance information about the soil vertical dimension. In addition to rapid and relatively inexpensive estimates of soil properties and processes, PSS can also rapidly provide information about the short-scale spatial heterogeneity of soils, which is of particular use in modeling soils (Kruger et al., 2013). Proximal soil sensing can also play a gap-filling role in increasing the level of spatial detail available from existing monitoring networks (Ochsner et al., 2013; Schirrmann et al., 2013), which will be important for soil process modeling that incorporates spatial processes.

As shown above, a number of areas of development exist that will improve the potential of PSS for soil process modeling. To realize this potential, the following objectives must be achieved: (i) automated interpretation of sensor data, using standardized calibration data sets and generally applicable calibration techniques, (ii) elimination of field- or sensor-specific effects on sensor data to allow calibration from a wide range of available data and sensor types,
(iii) multisensor or multiparameter readings to allow “snapshots” of all soil parameters of interest across the whole profile, and (iv) development of methods to allow cheap, mass-produced sensor devices (e.g., mobile-phone cameras) to be used in crowd-sourced information acquisition.

For each of the above objectives, significant progress has been made in recent years and will continue. In its current state, PSS can and does already benefit soil process modeling, and it is anticipated that future developments will increase the rapidity and ease with which data required for soil process model development, initialization, and validation can be acquired. The IUSS Working Group on Proximal Soil Sensing (http://www.proximalsoilsensing.org/, accessed 27 Apr. 2016) provides information and links to events and resources of relevance and is the forum in which developments in this area are discussed and disseminated.

**Soil Databases**

Soil information is the key to evaluating ecosystem services like water regulation, water retention, nutrient regulation, waste treatment, and food production (de Groot et al., 2002). With the help of computer-based geographic systems, many groups have generated geographical databases to organize and harmonize the huge amount of soil information generated during the last century. Soil databases enable the application of soil models at regional to global extents. Many national agencies around the world have organized their soil surveys in databases including SSURGO (Soil Survey Staff, 1995), with soil information mainly from the USA, the Australian Soil Resource Information System (Johnston et al., 2003), the National Soil Inventory of Scotland (Lilly et al., 2010), and the Soil-Geographic Database of Russia (Shoba et al., 2010).

Besides national databases, global efforts are underway to compile databases from different countries or generate new soil information through the implementation of multinational projects. These include the Soil and Terrain Database (SOTER; van Engelen and Ting-Tiang, 1995), at 1:5000000 scale, containing digitized map units and their attributes; the World Inventory of Soil Emission Potentials, WISE, (Batjes, 2009), from 149 countries; the Harmonized World Soil Database (Nachtergaele et al., 2008) and the Land Use and Cover Area Frame Survey from the European Union (http://eusoils.jrc.ec.europa.eu/projects/Lucas/, accessed 27 Apr. 2016; Toth et al., 2013). All these efforts manifest the need to organize and distribute soil information within the soil scientific community, and to make it available for interdisciplinary studies.

In 2006, the GlobalSoilMap, a global consortium that aims to create a digital map of the world’s key soil properties (Arrouays et al., 2014), was established. This global effort will provide access to the best available map of soil properties across the globe at a resolution of 3 arc sec (~100 m) along with its 90% confidence of prediction, in a consistent format at the depth ranges of 0 to 5, 5 to 15, 15 to 30, 30 to 60, 60 to 100, and 100 to 200 cm. The methods used for GlobalSoilMap consider the nature, availability, and density of existing soil data. For example, an initial approach to mapping soil carbon in the United States is based on a 1:250,000 soil map from the USDA-NRCS, in which the soil polygons were converted to raster estimates of organic carbon content for the six depth intervals of the GlobalSoilMap specifications (Odgers et al., 2012). Thus far, the most comprehensive example of soil property maps made according to GlobalSoilMap specifications is the Australian Soil and Landscape Grid (http://www.clw.csiro.au/aclep/soilandlandscapegrid/, accessed 27 Apr. 2016; Grundy et al., 2015). Other examples include the mapping of soil texture and organic carbon in Denmark (Adhikari et al., 2014). Another initiative is the Soilgrids by ISRIC (http://soilgrids1km.isric.org/, accessed 27 Apr. 2016), which used the GlobalSoilMap specifications except that the spatial resolution is 1 km (Hengl et al., 2014). The aforementioned databases in combination with pedotransfer functions (see the following section) have been successfully used to evaluate the impact of agricultural expansion (Maeda et al., 2010), global agricultural suitability (Zabel et al., 2014), nutrient stoichiometry under native vegetation groups (Bui and Henderson, 2013), and soil erodibility estimates (Panagos et al., 2012). In addition, global soil information should better inform global climate models (Wilson and Henderson-Sellers, 1985), hydrology models (Weiland et al., 2010), and road planning (Laurnance and Balmford, 2013).

**Informing Soil Models Using Pedotransfer Functions**

Pedotransfer functions, empirical relationships between parameters of soil models, and more easily obtainable data on soil properties have become indispensable 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 estimate parameters of soil models. Pedotransfer functions are extensively used in soil models addressing the most pressing environmental issues, such as carbon sequestration and gas emission, climate change and extreme events like floods and droughts, and soil ecological services and sustainability (e.g., Decharme et al., 2011; Piedallu et al., 2011; Wiesmeier et al., 2012). Currently, PTFs are mostly applied to estimating the soil water retention curve and soil hydraulic conductivity curve (Vereecken et al., 2010), solute transport parameters (Koestel et al., 2012), erosion and overland transport (Guber et al., 2014), and adsorption isotherms (Kodesova et al., 2011). However, the pedotransfer concept can be applied to any soil attribute. In particular, as the interest in modeling biogeochemical processes increases, development of PTFs for parameters of those processes will become essential. The process of PTF development is outlined in Fig. 6.

Because the equations to express PTF relationships are essentially unknown, a trend has emerged to employ machine-learning methodology (e.g., artificial neural networks, support vector machines, decision trees), which in theory is flexible enough to simulate highly
nonlinear dependences hidden in analyzed data. This methodology, however, comes with the penalty of a large number of coefficients that are difficult to estimate reliably. Applying a preliminary classification to PTF inputs and PTF development for each of the resulting groups holds the promise of providing simple, transparent, and more reliable pedotransfer equations. The existence of PTFs reflects the outcome of some soil processes; thus, using models of those processes to generate PTFs, or at least physics-based functional forms for PTFs, is an expected research avenue.

Pedotransfer functions are evaluated by their accuracy (i.e., errors with the development data set), their reliability (i.e., errors with data that have not been used in the PTF development), and their utility (i.e., errors of soil model where PTF-predicted parameters are used). Depending on the sensitivity of the soil model to PTF-estimated parameters, various levels of PTF accuracy and/or reliability may be acceptable in terms of the PTF utility (Chirico et al., 2010). The multiplicity of models (i.e., presence of several models producing the same output variables) is a typical feature in the PTF research field. However, PTF intercomparisons are lagging behind PTF development, aggravated by the fact that coefficients of PTF based on machine-learning methods are usually not reported. There is a pressing need to develop and implement protocols for PTF utility evaluation and intercomparison.

Estimating the variability of soil model parameters becomes increasingly important as newer modeling technologies (e.g., data assimilation, ensemble modeling, and model abstraction) become progressively more popular (Guber et al., 2006; Pan et al., 2012). The variability of PTFs rely on the spatiotemporal dynamics of soil variables, which open new sources of PTF inputs stemming from technology advances, such as monitoring networks, remote and proximal sensing, and omics (e.g., Tranter et al., 2008; Jana and Mohanty, 2011).

Burgeoning PTF development has so far not filled several persisting regional knowledge gaps. Remarkably little effort so far has been put 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. The challenge is to correct this situation in the near future. Soils from tropical regions are quite often considered as a pseudo-entity for which a single PTF can be applied (Minasny and Hartemink, 2011). This assumption will no longer be valid as more regional data are accumulated and analyzed. Other advances in regional PTFs will be possible because of the presence of large databases on region-specific useful PTF inputs such as moisture equivalent (Ottoni et al., 2014), laser diffractometry data (Lamorski et al., 2014), or soil specific surface (Khlosi et al., 2013).

Most transport models in soils—whether water, solutes, gas, or heat—involves parameters that are scale dependent. Recently, the need to match the scale of computational grid cells and scale of the flux parameter PTF estimation was shown (Pachepsky et al., 2014). Knowledge about scale effect on parameters is rapidly expanding for overland flow and transport (Delmas et al., 2012). Including scale dependencies in PTFs is the grand challenge in improving PTF usability.

Another scale-related challenge is PTF development for coarse-scale soil modeling, such as for land-use change or climate models. Soil parameters in these models cannot be measured, and the efficiency of PTFs can be evaluated only in terms of their utility (Gutmann and Small, 2007; Shen et al., 2014). There is an urgent need to determine combinations of pedotransfers and upscaling procedures that can lead to the derivation of suitable coarse-scale soil model parameters. Also, the coarse spatial scale often assumes a coarse temporal support, which requires an understanding of how to include in PTFs other environmental variables, such as weather and management attributes.

Temporal and spatial aspects of PTF development and applications have not received proper attention (Romano, 2004). Because PTF input variables demonstrate dependencies of spatial location and time, an effort will be made to determine whether PTF-estimated parameters have the same spatial and temporal correlations as measured ones, and whether regionalization and upscaling of PTF-estimated and measured soil parameters produce similar results. More efficient use of topography as an essential spatial covariate is also expected.

Pedotransfer functions are empirical relationships, and their accuracy outside the database used for 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 PTFs is to assist in modeling for screening and comparative purposes, establish ranges and/or probability distributions of model parameters, and create realistic synthetic soil data sets and scenarios. Further exploration is needed before using PTFs 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. Developing and improving PTFs will remain the mainstream way of packaging data and knowledge for applications of soil modeling.

Parameterizing Models with Nondestructive and High Resolution Water Stable Isotope Data

Physically based numerical soil–vegetation–atmosphere transfer models (SVAT) gather state-of-the-art knowledge on processes involved in the transfer of heat and water within the soil profile, on soil–plant relations (root water uptake and/or hydraulic redistribution), and on soil and plant–atmosphere interactions (radiative transfers and exchange of fluxes of momentum, heat and water vapor, i.e., ET). They are complex models that require careful calibration of their many parameters, which can be done by feeding them with high resolution input data, such as the temporal development of soil water isotopologue profiles.

For decades now, stable isotopologues of water ($^1$H$^2$H$^16$O and $^1$H$^2$$^18$O) have been used in identifying and quantifying sources and sinks as well as partitioning processes of terrestrial water, and hence are an invaluable source of information for improving soil hydrological and SVAT models. Mass differences of these heavy isotopologues relative to the most abundant water molecule ($^1$H$_2$$^18$O) lead to thermodynamic and kinetic isotopic effects, causing detectable differences in the isotopic composition ($^8$D$^2$H and $^8$D$^18$O) of water in different compartments, such as groundwater, surface water, soil and plant water, and atmospheric water vapor. These differences have been used to study groundwater recharge, atmospheric moisture circulation, water-balance closure of lakes, and reconstruction of root water uptake profiles, as well as for ET partitioning from the plot to the global scale, (e.g., Craig, 1961; Moreira et al., 1997; Yakir and Sternberg, 2000; Gibson, 2002; Williams et al., 2004; Nippert et al., 2010; Rothfuss et al., 2010; Wang et al., 2010; Jasechko et al., 2013).

The first analytical description of water isotopologue profiles for an isothermal and saturated soil at steady state was proposed by Zimmermann et al. (1967), which was later extended to nonsaturated profiles under non-steady state and nonisothermal conditions (Allison et al., 1983; Barnes and Allison, 1983; Barnes and Allison, 1984; Barnes and Walker, 1989). These analytical formulations link the shape of the water isotopologue profiles to soil evaporation flux and regime and to the soil physical properties associated with both diffusive and convective water transport (such as tortuosity length and dispersivity). In soils between rain events, the combined action of convective capillary rise of water depleted in the heavy stable isotopologues with back-diffusion of water enriched in the heavy stable isotopologues from the evaporation site (i.e., soil surface or evaporation front) downward leads to the formation of (typically exponential) soil water stable isotopologue profiles.

More recently, the movement of $^1$H$^2$H$^16$O and $^1$H$^2$$^18$O was implemented in various SVAT models, including TOUGHREACT, SiSPAT-Isotope, Soil–Litter–Iso, and HYDRUS 1D (Singleton et al., 2004; Braud et al., 2005; Haverd and Cuntz, 2010; Rothfuss et al., 2012; Sutanto et al., 2012). In addition to the mass conservation equation for water, these models solve an equivalent conservation equation for the water isotopologues $^1$H$^2$H$^16$O and $^1$H$^2$$^18$O and need isotopic initial and boundary conditions. Fluxes of water isotopologues are considered throughout the entire soil profile, that is, in both vapor and liquid phases, and not only in the vapor phase above a so-called evaporation front (the minimal depth where nonequilibrium gas exchange occurs in the soil (defined as the minimal depth where nonequilibrium gas exchange occurs in the soil, Rothfuss et al., 2015), or only in the liquid phase below it. In addition, and contrary to, for example, the study of Barnes and Walker (1989), these numerical models do not make use of a similarity variable, proportional to depth and (time)$^{-1/2}$, and do not require particular boundary conditions for the computation of $^1$H$^2$H$^16$O and $^1$H$^2$$^18$O profiles. In addition to thermodynamic (equilibrium) isotope effects, which are only temperature dependent, kinetic isotope effects during soil evaporation greatly affect the stable isotopic composition of soil water and evaporation and can be highly variable (Braud et al., 2009).

Thus, a better understanding of the implications of these kinetic effects in addition to the well-characterized equilibrium effects and their implementation in SVAT models are required for improving the use of $^1$H$^2$H$^16$O and $^1$H$^2$$^18$O as tracers of soil water processes. An important challenge is to provide those models with high-resolution isotope data, both in space and time. Moreover, parallel to field studies, effort should be made to design specific experiments under controlled conditions, allowing underlying hypotheses of the above-mentioned isotope-enabled SVAT models to be tested. Using isotope data obtained from these controlled experiments will improve the characterization of evaporation processes within the soil profile and ameliorate the parametrization of the respective isotope modules.

Soil water $^8$D$^2$H and $^8$D$^18$O typically have been measured by destructive sampling, followed by cryogenic soil water extraction (e.g., Aragüés-Argüás et al., 1995) and offline analysis with isotope-ratio mass spectrometers. Although this time-consuming and labor-intensive procedure provides high-quality data, it has only poor temporal and spatial resolution. As a consequence, measurements of the isotopic composition of evaporation, inferred from that of soil water at the evaporative site in the soil, are still sparse, but crucial to constraining transpiration over ET ratios, (e.g., Dubbert et al., 2013; Hu et al., 2014). Another challenge is therefore to develop new methodologies toward monitoring soil water $^8$D$^2$H and $^8$D$^18$O online with high resolution and in a nondestructive
manner. The first successful attempt was made using microporous polypropylene tubing combined with laser-based infrared spectrometers (Rothfuss et al., 2013, 2015; Volkmann and Weiler, 2014). These methodologies have also been applied to both laboratory and field experiments and compared with traditional methods (e.g., cryogenic distillation) for determining soil-water $^2$H and $^{18}$O signatures (Gaj et al., 2015; Gangi et al., 2015). Another exciting challenge of the coming years is to determine plant-root water-uptake profiles via online and nondestructive determination of soil water $^2$H and $^{18}$O profiles using microporous tubing or membrane-based setups. These high resolution nondestructive isotope data will drastically improve the basis for constraining the above-mentioned SVATs through, for example, inverse modeling and within the framework of specific (controlled conditions) experiments.

Toward a Soil Modeling Platform

Since the advent of computer technologies in the 1980s, we have seen an unprecedented development of mathematical models that are able to simulate soil processes at an ever-increasing complexity and at scales ranging from the pore to continents. Many of these efforts have been undertaken by specific soil science disciplines or communities focusing on specific processes and scales, leading to a diverse landscape of soil models. In this section we will discuss recent developments that aim to better integrate and improve exchange of knowledge, such as the establishment of a virtual soil modeling platform, the development of technologies to couple models, the establishment of benchmark initiatives, and soil modeling intercomparison studies. Finally, the soil modeling community should reach out to other communities that explicitly deal with soil either as an environmental compartment controlling key ecological, climatic and hydrological processes or as the substrate for producing crops and biomass. A recent initiative, the International Soil Modeling Consortium (ISMC; https://soil-modeling.org/, accessed 27 Apr. 2016), has been established as a community effort to address the current challenges of soil modeling.

Virtual Soil Platform

In the environmental and soil science communities, the need for coupling models and the associated knowledge has only emerged recently. The development of a coupling tool or modeling platform is mainly driven by the necessity to create models that consider multiple processes and that take into account feedbacks between these processes. Soil models often focus on specific processes, compartments, and scales, and they are often developed for specific applications. The development of a modeling platform may constitute an efficient and rapid way, not only to address emerging challenges, such as predicting soil functions and soil evolution under global change, but also to share our vision on soil functioning at different scales and to strengthen collaboration among soil scientists, soil modelers, and the Earth-system research community. Such a modeling platform goes beyond the coupling tools that have already been proposed, including OMS3 (David et al., 2013), CSDMS framework (Peckham et al., 2013), and the Open MI project developed within the framework of the European Community (http://www.openmi.org/, accessed 25 Apr. 2016). We should expect a modeling platform that is more ambitious than the coupling of existing numerical codes and one that shares underlying principles and knowledge. We need to develop complex models that enable us with tools that bring responses to current issues on soil functioning and soil evolution within the framework of global change. We also need to share in a common framework our visions of soil functioning at various scales—to both strengthen our collaborations and to make them visible to other communities working on environmental issues.

We therefore propose to develop a virtual soil platform (VSP) that serves as a hub for sharing soil process knowledge, modules (i.e., numerical tools and algorithms simulating a process), and models (i.e., a logical combination of several modules) and that addresses the issues discussed above. Virtual soil platform should enable soil scientists not familiar with model development to develop numerical representations of soil processes or to build their own models. To make this possible, VSP should enable an easy exchange of processes, variables, modules, and models between users. The VSP should provide access to tools enabling sensitivity studies, parameter estimation, stochastic analysis and ensemble runs, data assimilation, visualization of simulation results, and model comparison and benchmarking (recall section “Does Local-Scale Model Complexity Matter for Predictions at Larger Scales?”). In addition, VSP should be linked to soil databases providing information on soil properties, spatial variability (see “Novel Optimization Methods and Their Application to Soil Modeling”), boundary conditions, validation data sets, and so on. The purpose is to offer a common tool facilitating not only the exchange of knowledge, the reuse of recognized modules and models and the development of new ones, and the access to various peripheral tools, but also the exchanges between users.

At present, the VSoil platform (Lafolie et al., 2015) is being developed (http://www6.inra.fr/vsoil, accessed 27 Apr. 2016). It addresses the issues listed above and may serve as a starting point toward the future development of the ISMC. More precisely, VSoil offers a means of dealing with processes, not just with codes representing these processes. Processes are clearly defined. This means that all the entities (i.e., states, parameters, constants and fluxes) that describe processes and all the output VSoil produces are listed and visible to anyone using the framework, without having to access the codes of the modules. The processes and entity lists are open, as new items will be progressively added. VSoil clearly differentiates between process knowledge, the various mathematical representations of soil processes, and their numerical...
implementation, thus favoring the use of the framework by those not familiar with modeling. By using sets of processes and variables, VSoil automatically ensures that the connections between processes and modules are checked for compatibility when assembled for constructing a model. Having a set of uniquely defined entities (i.e., definition and units) on which models can draw is also essential, given that a reasonable objective is to couple the platform with databases for model comparison, data assimilation, variables forcing, or parameter estimation. In addition, a well-defined set of variables is fundamental when collaboration between people from various fields of expertise (physics, biology, chemistry, and so on) is sought. We view this as a goal for tools dedicated to the development of complex soil functioning models. Thus, we suggest that effort be focused on the sharing of knowledge in addition to all that can be accomplished in sharing and coupling numerical tools.

VSoil eliminates all the portability (compilation, version, and so on) problems that arise when exchanging computational tools. In addition, given that the platform manipulates processes and variables, and that modules are linked to a process, all information about a module or model is readily visible and not hidden somewhere in the code. In particular, the lists of exchanged variables are explicitly displayed, as well as the list of parameters for a module. Using a platform based on processes and modules also eases collaboration between coworkers since agreement on concepts and variables can first be reached. Numerical code development can be performed after this stage; this phase can be split into several tasks that can be realized simultaneously in different places if needed, without worrying about compatibility or portability. Hence, working within a common framework would intensify communications and exchanges, speed up model development, promote the reuse of well-recognized tools, and offer visibility to models developed by the soil science community.

**Model Coupling Approaches**

In complex systems, such as soils, mathematical models generally describe several distinct but simultaneously occurring processes. The full mathematical model can often be split into several distinct modules; a solution of the full model is achieved by operator splitting techniques. Or, in a bottom-up view, several models describing distinct processes can be coupled together to characterize a more complex system. In this way, additional processes can be integrated as new modules if required for a specific scientific problem. This approach also allows an exchange of modules, which enables the user to analyze the impact of different modeling approaches.

Coupling methods include (i) light coupling that is based on shared input/output files, (ii) external approaches with a central coupler, or (iii) full coupling, using integrated classes or subroutines. The advantage of the light-coupling approach is that models are independent executables and only need to share the same format for the input/output files. One example of this approach is the coupling of SOILCO2 and RothC (Herbst et al., 2008b) where the CO2 production rate required by SOILCO2 is computed by the RothC model. Another example is the coupling of the dynamic root architectural model RootBox with the model for water flow in soil and root system, R-SWMS (Leitner et al., 2014). Here, RootBox computes the geometry of the growing root system used by R-SWMS. The disadvantage is that it is relatively inefficient compared to other approaches.

A minimally intrusive coupling approach attaches independent models to a central coupler such as OASIS (https://verc.enes.org/oasis, accessed 27 Apr. 2016) or MCT (http://www.mcs.anl.gov/research/projects/mct/, accessed 27 Apr. 2016). Here, each model must include a piece of software that enables communication with the central coupler; thus, a slight change to the code is necessary. The coupler establishes the global communication and memory space; it exchanges data in memory instead of time-consuming I/O procedures. A further advantage of this approach is that it facilitates the running of models not only individually but also while in ensemble (for data assimilation) or Monte Carlo mode (for uncertainty analysis), as well as the coupling of further computational tools, such as inversion algorithms for parameter estimation. Examples of this approach are more commonly found in the Earth system community (Warner, 2008).

**Benchmarks and Soil Model Intercomparisons**

Model verification, benchmarking, and intercomparisons are activities that are intrinsically linked with the development of complex mathematical models simulating various processes in soils. Because of the inherent heterogeneity of soil environments, the temporal and spatial variability of boundary conditions, and the nonlinearity processes and various constitutive functions, general solutions of the governing mathematical equations are usually achieved using numerical approximations. Given the diversity of processes and numerical approaches, scientists and model developers must verify and test their models or demonstrate their models were independently verified and tested. Verification of a code should ensure that the equations constituting the mathematical model are correctly encoded and solved. Verification of a code consists of showing that the results generated by the model for simpler problems are consistent with available analytical solutions or are the same as, or similar to, results generated with other numerical codes (model intercomparisons). The latter procedure is also called benchmarking.

Available analytical solutions are often limited to idealized transport domains, homogeneous and isotropic media, and uniform initial and constant boundary conditions. The very reason for developing numerical models is to go beyond the range of available analytical solutions (i.e., to allow irregular transport domains, heterogeneous and anisotropic media, variable boundary conditions, and nonlinear processes). Verification in such conditions is often
accomplished using model intercomparisons that use approximate tests for internal consistency and accuracy, such as mass conservation, global mass-balance errors, and sensitivity to changes in mesh size and time steps.

In the literature, many model intercomparison studies have been reported for subsurface flow and transport models. For example, Scanlon et al. (2002) compared water-balance simulation results from seven different codes (HELP, HYDRUS-1D, SHAW, SoilCover, SWIM, UNSAT-H, and VS2DTI) using 3-yr water-balance monitoring data from nonvegetated engineered covers (3-m deep) in warm (Texas) and cold (Idaho) desert regions. Vanderborght et al. (2005) developed and used a set of analytical benchmarks (of differing complexity) to test numerical models (HYDRUS-1D, MACRO, MARTHE, SWAP, and WAVE) of flow and transport in soils. Oster et al. (2012) compared the simulated crop yields grown under production practices and transient conditions (involving pressure head and osmotic stresses) in the western San Joaquin Valley of California, using the ENVIRONMENTAL HYDRUS-1D, SALTME, SWAP, and UNSATCHEM models. Finally, intercomparisons of results obtained by PEARL, PELMO, PRZM, and MACRO models for nine (MACRO only for one) FOCUS scenarios/sites, which collectively represent agriculture (and different climate regions) in the EU, for the purposes of a Tier 1 EU-level assessment of the leaching potential of active substances were performed by the FOCUS group (FOCUS, 2000).

Similar efforts are being performed in related environmental fields. For example, Hanson et al. (2004) evaluated 13 models varying in their spatial, mechanistic, and temporal complexity for their ability to capture intra- and interannual components of the water and carbon cycle for an upland, oak-dominated forest of eastern Tennessee. A set of well-described benchmark problems that can be used to demonstrate model conformance with norms established by the subsurface science and engineering community has recently been developed for complex reactive transport numerical models (CrunchFlow, HP1, MIN3P, PFlotran, and TOUGHREACT) (e.g., Rosenzweig et al., 2013; Steefel et al., 2015; Xie et al., 2015). Rosenzweig et al. (2013) described the Agricultural Model Intercomparison and Improvement Project (AgMIP), which is a major international effort linking the climate, crop, and economic modeling communities with cutting-edge information technology to produce improved crop and economic models and the next generation of climate-impact projections for the agricultural sector. Finally, the World Climate Research Program (WCRP) Working Group on Coupled Modeling (http://www.wcrp-climate.org/wgcm/projects.shtml, accessed 27 Apr. 2016) catalogs a large number of Model Intercomparison Projects (MIPs) related to various climate-related models.

Similar model intercomparison studies will undoubtedly continue as advances in measurement technology, computing technology, and numerical techniques enable the development of models of ever-increasing levels of sophistication that cannot be readily verified using analytical solutions such as those developed and/or suggested by Vanderborght et al. (2005). The soil-modeling community should thus expand on this work by establishing a benchmark and validation platform with standardized and high-quality data sets that would use common data formats, protocols, and ontologies and that would be readily available to model developers for further model testing and intercomparisons. Ontologies refer to a standardized vocabulary enabling a common understanding of the exact meaning of different terms (e.g., parameters, variables) used in a science community. Examples can be found in biology (http://www.plantontology.org/, accessed 27 Apr. 2016) or agriculture (http://aims.fao.org/vest-registry, accessed 27 Apr. 2016). The database could include not only experimental data sets, but also input and output files of most commonly used soil models applied to these data sets.

**Linking Soil Modeling Platforms with Climate, Ecology, and Hydrology**

It is clear that soil plays a vital and pivotal role in environmental responses to climate change and variability, in ecological vigor and hydrologic extremes, and in the outcomes of models used to understand the strength and direction of these connections. Many of these models focus on the supporting processes of soils, particularly related to water cycling (stocks and fluxes of water into/from the soil profile) and nutrient (C, N, P) cycling, which are closely linked to provisioning services. The models also simulate regulating services, described by Dominati (2013) as flood mitigation, filtering of wastewater, and so on. Predictive and hindcast models used across scientific disciplines can provide substantial insights into ecosystem processes and services, as well as into the intricate connection among the different pools of natural resources provided by soil.

As described by Sellers et al. (1997), land–atmosphere models have evolved into sophisticated soil–vegetation–atmosphere systems that provide large-scale transfer of water vapor and carbon. Many aspects of these climate circulation models connect to surface processes and the uppermost soil horizons of land. These processes involve understanding soil hydrology, impacts on the soil’s energy balance, and ecological response to climate and climate variability, all of which impact soil properties, formation, and processes that influence soil formation and degradation. This knowledge base is being implemented, although slowly and at variable spatial and temporal scales, into numerical codes that simulate biospheric processes.

We see the effective incorporation of these provisioning and regulating processes into scale-appropriate models as a significant challenge, and one that could expand soil modeling applications to other scientific disciplines. For example, Ochsner et al. (2013) discussed the connection of soil water storage and content
to ecological function, biogeochemistry, and ecological model platforms. This vital link between soil and ecosystem services is parameterized by lumping many soil processes into compartments in which reactions occur. The CENTURY/DAYCENT model (Parton et al., 1998) focuses on carbon and nutrient dynamics, and biosphere models like SiB (Sellers et al., 1986) and BATS (Dickinson et al., 1986) simulate SVAT. These and other models are now being widely used by the ecological and biogeochemical communities, even though they generally do not use physically based governing equations or constitutive relations when incorporating soil processes; the soil modeling community can make highly relevant contributions in this regard. For example, recently Ren et al. (2008) explicitly accounted for vegetation canopy and physiological control of ET and soil water budgets, improving water budget estimates deeper into the soil profile rather than matching soil response for the upper (15-cm) soil layer only.

Hydrologic models have for some time generally included soil property parameters, though to varying degrees of sophistication. Regulating water exchange and movement are critical for accurately predicting soil (and deeper) recharge, surface-runoff timing and severity, and the ET component of hydrologic models that ultimately connect to climate or atmospheric codes. One-dimensional approaches (e.g., HYDRUS-1D; Šimůnek et al., 2008) are used extensively in the agricultural and environmental fields. These often solve the Richards equation under variably saturated conditions, using common forms of soil water retention and hydraulic conductivity curves. But, these approaches are less commonly used in landscape-scale approaches for water routing like SWAT (Arnold and Fohrer, 2005; Chen et al., 2011) or HSPF (e.g., Donigian et al., 1995; Bicknell et al., 1997), which rely on a “bucket model” approach and the concept of field capacity and gravitational downward flow. There remains a divide between physically based models at small spatial and temporal scales and lumped parameter models for landscape-type applications. This divide exists because of computational limitations (i.e., lack of sufficient memory or high-performance computer resources) or theoretical limitations. In the latter case, soil physicists do not deem pore-scale approaches scalable to landscape and regional scales. Bridging this divide and using manageable soil properties and governing equations across scales is a significant challenge that needs to be overcome for hydrologic models to be useful for decision makers.

Increasingly, integrated modeling platforms are collaboratively developed, with model advancements occurring through specific modules that spread scientific expertise across disciplinary boundaries. An excellent example is the Community Earth System Model (CESM), maintained by the National Center for Atmospheric Research (NCAR). Among the principle modules of this global model is the Community Land Model (CLM), the purpose of which is to improve understanding of natural and human impacts on vegetation and climate at the regional or global scales. The CLM includes surface heterogeneities and consists of submodels that represent the hydrologic cycle, biogeochemical cycling, and ecosystem dynamics (Lawrence et al., 2011), many of which fit neatly into the framework of Dominati (2013) that connects soil capital to ecosystem processes and services. The CLM is well suited to study the role of land processes in weather and climate change, and efforts are being devoted to improve the representation of the role of subsurface processes. For example, the mechanistic ParFlow model was recently coupled to the CLM (Kollet and Maxwell, 2008; Maxwell, 2013) for regional-scale applications, with the ability to simulate complex topographies, geology, and subsurface heterogeneities of the coupled vadose zone–groundwater system. A challenge for the modeling community would be to incorporate nutrient cycling, erosion, and other supporting–degrading processes at spatial and temporal scales that can facilitate the tracking of ecosystem services through time by changing land use and climate.

For the future, a persistent question is how to effectively incorporate soil properties, taken at the point scale, into larger-scale (i.e., landscape, watershed) models that simulate ecological, biochemical, and climatological (supporting) processes. The SoilML standard for soil data transfer and storage (Montanarella et al., 2010) may help in this process. Moreover, whereas advanced soil modeling platforms increasingly integrate physical, chemical, and biological processes that couple climate with hydrology and geochemistry, much of the biological components remain relatively underdeveloped. In part, much of the microbiological system remains a black box in many soil based models, especially as related to microbial kinetics and effects of the dynamics of soil environmental changes (water, temperature, nutrients) on microbial processes. Although much experimental work is being done to understand soil fauna (e.g., fungi, worms) and how they alter the soil environment, we are unaware of soil modeling work that incorporates soil fauna impacts on the soil–climate system. Finally, because the main purpose of the IPCC and Millennium Development Goals (MDG) is to provide science for policy, and given the ongoing interest in incorporating ecosystem services into sustainable land management decisions, soil modeling platforms need to be designed to more effectively integrate soil modeling output into policy decisions at the regional and global scales.

Linking Soil Modeling Platforms with Crop and Biomass Production

Biomass production as an ecosystem service (“Biomass Production for Food, Fiber, and Energy” section) is strongly dependent on soil and crop interactions. Crop growth and development, as well as yield formation, are complex processes with dominant anthropogenic influence. Besides the genetic characteristics of crop species and crop cultivars, atmospheric conditions, soil properties and soil processes, crop growth depends on the intensity of crop and soil management. In general, in intensive high-input cropping systems
under irrigation, the farmer is able to optimize management in a way that the growth of a specific crop is only constrained by radiation and air temperature (potential production conditions). However, in terms of area, irrigated cropping systems have a low share in the global cropland, and rainfed systems are predominant, where, depending on climate and soil water retention curve, soil water is a major constraint. Therefore, among existing dynamic crop models, the majority considers the soils’ function in storing infiltrated water and supplying it to the crop. However, the level of detail of the representation of this important soil function and its interaction with crop roots and crop water demand is highly variable. Most crop models use a one-dimensional conceptual approach such as a bucket type approach to characterize the dynamics of soil water storage, either in a one layer or in a multilayered soil (DSSAT, Jones et al., 2003; APSIM, Keating et al., 2003; MONICA, Nendel et al., 2011). Physically based approaches to simulate soil water fluxes integrated into dynamic crop models are rather scarce (DAISY, Abrahamsen and Hansen 2000). The SIMPLACE platform offers three different one-dimensional approaches to simulate soil water dynamics, which can be combined with two different approaches of root development and three different crop water uptake mechanisms (Gaiser et al., 2013). Depending on the availability of input data, prevailing water management practices, and the climatic conditions where the model is to be applied, simple or more complex combinations can be selected by the user.

To be suitable for cropping systems with reduced management intensity, crop models must consider additional soil processes that are related to crop nutrient supply and in particular to nitrogen. However, due to the fact that soil nitrogen dynamics including mineralization and immobilization, leaching, nitrification, denitrification, volatilization, and crop uptake are extremely complex, different approaches with varying levels of detail have been implemented or coupled with crop growth processes. In cropping systems where application rates of mineral nitrogen fertilizers are on the order of potential crop demand, only the uptake of the applied mineral N may be considered to cover the actual or daily crop N demand in the simulations. In organic agriculture or in low-input systems, such as in small-holder subsistence farms in developing countries, soil nitrogen routines must consider the nitrogen mineralization and immobilization processes linked to soil organic matter. Usually, the more complex soil nitrogen routines in existing crop models consider different soil nitrogen pools (linked to soil carbon pools), and their respective decomposition and mineralization rates are calculated taking into account environmental variables like soil moisture, soil temperature, or soil clay content (CENTURY, Parton et al., 1992; CANDY, Franko et al., 1995; DAISY, Abrahamsen and Hansen, 2000; SIMPLACE, Gaiser et al., 2013). Crop nitrogen uptake is then driven by the amount of soil mineral N over the rooted zone, crop nitrogen demand, and in some cases the density and N uptake capacity of the roots in the respective soil layers. Nitrogen leaching, as an important process in humid climates, is usually also considered in these more complex soil nitrogen subroutines, whereas other soil related processes like nitrification, denitrification, or ammonium fixation and volatilization are implemented in only a few models (e.g., DNDC, Kraus et al., 2015; and CropSyst, Stöckle et al. (2014)).

Besides nitrogen as one of the major crop nutrients, there are only a few crop models that consider phosphorus as a limiting factor with (CropSyst, APSIM, DSSAT, SIMPLACE, EPIC; Williams and Izaurralde, 2005) or without (WOFOST, van Ittersum et al., 2003; Lintul5, Lefelaar, 2012) taking into account the dynamics of adsorption or fixation of inorganic P onto the soil matrix or the transformation of organic soil P. Among other major crop nutrients like K, Mg, Ca, or S, only K is taken into account by four dynamic crop models either with (EPIC Version EPICSEAR; De Barros et al., 2004) or without (WOFOST, Lintul5, SIMPLACE) associated transformation and adsorption processes in the soil. To our knowledge modeling of the availability of micronutrients in the soil or their uptake by the crop is still a gap when coupling soil processes with crop and biomass production, although micronutrient deficiency is a well-known obstacle to advance intensification and increase yields on highly weathered soils in Africa, Asia, and South America (Voortman et al., 2003).

Modeling soil conditions that are adverse to crop growth (e.g., salt toxicity, water logging, soil compaction, aluminum, and iron toxicity) and quantification of their impacts on crop roots and crop growth are another bottleneck when coupling soil processes with crop and biomass production. The crop models EPIC and STICS use different relationships between either soil strength (Williams and Izaurralde, 2005) or soil bulk density (Brisson et al., 2003) and root elongation rate to describe the effect of soil compaction on root growth and, subsequently, water and nutrient uptake. In addition, the EPIC model estimates the effect of aluminum toxicity on root growth by relating Al saturation in the soil to a crop specific maximum Al saturation threshold (USDA, 1990). In a more recent Windows-based version of EPIC, the effect of increased soil electrical conductivity as a measure of high salt concentrations in the soil on crop growth had been incorporated (Gerik et al., 2013). Water logging can also be an important growth limitation to crops and in particular to roots. The processes leading to water logging, that is, permanent saturation of the root zone with water, can be manifold and involve reduced percolation of rainwater, occurrence of surface water flooding, and groundwater rise. One the one hand, modeling of water logging therefore requires detailed parametrization of soil hydraulic conductivity curve and reliable estimation of one-, two-, and three-dimensional soil water fluxes, including landscale-scale hydrological processes in the case of groundwater influence of flooding from adjacent surface water streams. On the other hand, modeling the crop-specific, physiological response of the root system and its interaction with the shoot is neither fully understood nor adequately implemented in crop models. A first attempt to cover some of the challenges was made recently by Stöckle et al. (2014) through coupling a landscape-scale hydrological model.
with CropSyst (CropSyst-Microbasin). In summary, there are many interfaces between soil processes, the crop roots, and their interaction with the shoot that ultimately determine crop yield and biomass production, all of which require further investigations at the plot, field, and landscape level and subsequent implementation into coupled soil–plant modeling platforms to simulate biomass production under a wide range of climate, soil, and management conditions.

Regarding the technical implementation of crop simulation models, besides a wide range of one package crop simulation models, there are several crop-simulation environments relying on a modular structure to describe crop-growth processes at the field scale. These environments all consider aboveground and belowground processes, but with different degrees of detail (Keating et al., 2003; Donatelli et al., 2010). Examples for developments in Germany are SIMPLACE (http://www.simplace.net, accessed 26 Apr. 2016; Gaiser et al., 2013), ENREF_170, Expert-N (https://soil-modeling.org/models/model-descriptions/expert-n), MONICA (Nendel et al., 2011), and HUME (Kage and Stuetzel, 1999; Ratjen, 2012). At the global scale, the DSSAT (http://dssat.net/, accessed 26 Apr. 2016) platform is quite wide-spread for one-dimensional applications from field to region. As an example for three-dimensional applications, the OpenAlea (Cokelaer et al., 2010) open-source project should be mentioned.

#### Summary and Outlook

Since the early attempts in systematic modeling of soil processes that emerged with advances in analog and digital computers in the midst of the 20th century, there has been great progress across a broad range of space and time scales (pores to catchments and seconds to decades). Yet, our current understanding of the complexity of soil processes and ability to observe these processes at ever-increasing resolution point to significant gaps in representing this critical compartment of biosphere. The growing importance of soil in a host of topics and its central role in a range of ecosystem services, climate, food security, and other global terrestrial processes make quantification and modeling of soil processes an urgent challenge for the soil science and neighboring communities. We focused on identifying various key challenges in modeling soil processes that are directly related to the hierarchical and complex organization of soils and soil systems and the functioning of soils in providing ecosystem services to society. Many of these challenges have been addressed in the individual sections, and here we identify four overarching grand challenges, summarized in Table 5, that we think will dominate in the field of soil processes modeling.

The first challenge is that of sharing knowledge across disciplines. It comprises the need to exchange knowledge about soil processes modeling across the different soil disciplines, and among other Earth, ecology, and plant sciences. Typically, many available soil models have been developed within different communities and disciplines addressing specific research questions covering a broad range of scales and often serving different purposes. Integrating our knowledge of soil process modeling in climate models, crop growth models, and ecological models may enhance our understanding of the complex interactions between the different compartments and their feedback mechanisms. The development and establishment of a community modeling platform could facilitate the exchange of knowledge on modeling soil processes, provide techniques and approaches to efficiently couple soil processes, and develop integrated models and benchmarks to test existing and newly developed models. The platform could also serve as a link with other disciplines listed above. A better interaction of the soil science community with other Earth science disciplines may enhance our understanding of soil processes in the landscape by, for example, coupling state-of-the-art approaches to soil infiltration with overland flow modeling, particle detachment, transport and deposition modeling across a heterogeneous landscape or by coupling of soil physical and chemical processes and soil biology to better understand and quantify supporting and degrading processes and key ecosystem services. The soil supporting and degrading processes and ecosystem services described here are determined by the combined effect of a multitude of individual processes. We are convinced that improved modeling of soil processes will also lead to a better quantification and prediction of ecosystem services. The development of more complex soil models and soil modeling platforms, together with the availability of novel experimental techniques, will also allow us to design new experimental setups based on soil model simulations, which will then enable the retrieval of soil properties that are difficult to measure.

<table>
<thead>
<tr>
<th>Table 5. Overarching challenges to modeling soil processes.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overarching challenges</strong></td>
</tr>
<tr>
<td>Challenge 1</td>
</tr>
<tr>
<td>Challenge 2</td>
</tr>
<tr>
<td>Challenge 3</td>
</tr>
<tr>
<td>Challenge 4</td>
</tr>
</tbody>
</table>
The second overarching challenge for soil modeling is the integration of pore- and local-scale soil process modeling into field-scale to global-scale land surface models, crop models, climate models, and terrestrial models of biogeochemical processes. These complex codes address issues such as parameterization of root water uptake processes, biotic processes, and upscaling of hydraulic properties, chemical and biological properties, and others. Effective integration will require the development of upscaling methods and approaches to derive effective parameters and equations that allow us to include pore- and local-scale process understanding, so we can describe processes at the field scale and beyond. Upscaling soil processes beyond the field scale will require us to embed and couple soils and soil process modeling into a landscape setting. This will entail a larger focus on nonlocal processes that are controlled by lateral water, energy, and matter fluxes. Lateral groundwater flow plays an important role in linking these processes because it influences, in part, the water table depth and its important consequences for soil water contents and water fluxes. Lateral fluxes in the atmosphere also play an important role for determining the upper boundary of the soil system. Besides lateral water and energy fluxes, lateral fluxes of soil material also become important when considering soil building processes over longer time scales. These processes need to be coupled with predominantly vertical fluxes of dissolved substances.

The third challenge embraces the monumental task of quantitative description of soil biotic processes at scales ranging from microbial activity at pores or on root surfaces to the emergence of vegetation patterns over extensive landscapes. At the core of this challenge is the representation of highly adaptive and dynamic biological processes that respond in new and surprising ways to changes in climate, land use, and management practices and their upscaling to represent fluxes and changes in soil properties at agronomic or climatic relevant scales. The rapid advances in remote observational methods and molecular genetic capabilities necessitate advanced modeling frameworks for effective integration of new observations with process understanding. Especially the upscaling of soil biotic processes may benefit from novel measurement techniques that enable to quantify and visualize microbial processes at the pore-scale level and at the interfaces of water and soil matrix. An important component of this is the need to agree on a framework of describing the soil microbial community in a manner that allows its functional dynamics and interaction with soil physical, chemical, and biological components to be described for modeling purposes, without oversimplification or loss of meaning.

Finally, we need to address the question of how to value ecosystem services using soil properties and processes in the proposed integrated modeling approach. We have used an ecosystem’s framework to identify the role and importance of soil modeling in characterizing and quantifying ecosystem services and we have identified specific challenges for improving soil process modeling. From a soil modeling perspective, we may want to challenge our soils community to work with ecologists, sociologists, and economists to develop such a framework that allows differentiation of soils based on their functioning properties and includes land use and/or tracking changes of supporting–degrading processes toward building spatial maps that quantify ecosystem services. This would be highly significant as the soil community’s contribution to local and regional policy and decision making and toward providing sustainable options for future land use and land use changes.

To meet these challenges, an international community effort is required, similar to initiatives in systems biology, hydrology, and climate and crop research. We are therefore establishing an international soil modeling consortium (https://soil-modeling.org/, accessed 26 Apr. 2016) with the following aims:

- establishing formal structures for guiding and building community-wide capabilities (repository, conferences, journal sections, liaisons with societies) to bring together experts in soil process modeling within all major soil disciplines;
- addressing major scientific gaps in describing key processes and their long-term impacts with respect to the different functions and ecosystem services provided by soil;
- intercomparing soil model performance based on standardized and harmonized data set;
- providing adaptive and peer-reviewed protocols for model components, benchmarking and testing, input information, ontologies, and data formats;
- integrating soil modeling expertise and state of the art knowledge on soil processes in climate, land surface, ecological, crop, and contaminant models;
- linking process models with new observations, measurements, and data- evaluation technologies for mapping and characterizing soil properties across scales;
- developing partnerships with similar modeling endeavors, industry, and funding agencies.

The consortium will bring together modelers and experimental soil scientists at the forefront of new technologies and approaches to characterize soils. By addressing these aims, the consortium will improve the role of soil modeling as a knowledge dissemination instrument in addressing key global issues, and we will stimulate the development of translational research activities.

References


Vegetation Science | Advancing Critical Zone Science | p. 41 of 57


Lefelaar, P. 2012. LINTUL-S crop growth simulation model for potential, water-limited, N-limited and NPK-limited conditions. Plant Production Systems Group, Wageningen University and Research Center, Wageningen, the Netherlands.


