

DEVELOPMENTS IN  
EARTH SURFACE PROCESSES

19

# PRINCIPLES AND DYNAMICS OF THE CRITICAL ZONE

JOHN R. GIARDINO AND CHRIS HOUSER



SERIES EDITOR: J. F. SHRODER JR

## **Chapter 4**

# **Regolith and Weathering (Rock Decay) in the Critical Zone**

Gregory A. Pope Department of Earth and Environmental Studies, Montclair State University, Montclair, New Jersey, USA

### **Abstract**

Weathering, or rock decay processes are at the center of the Critical Zone. The weathering engine modifies the Earth's crust to adjust to surface atmospheric, hydrologic, and biotic conditions, relevant to the many fields that find interest in the Critical Zone. Regolith refers to the body of decayed rock or sediment, coincidental to (and parallel and synonymous with) the Critical Zone. Weathering processes are synergistic, involving combinations of mechanical and chemical decay, taking place mostly at a nano-scale boundary layer at the mineral surface. Numerous factors intrinsic to the rock itself and of the surrounding subsurface and subaerial environment influence the availability of weathering agents and the rates and efficacy of decay. When environmental conditions allow, the terrestrial Critical Zone, regolith, can attain significant thickness, tens of meters, in most environments. Rates of formation calculate the oldest regolith profiles to exceed one million years in age.

**Keywords** weathering rock decay pedogenesis weathering system science regolith Critical Zone

### **4.1. Introduction**

Weathering and the Critical Zone have been inextricably linked, as both nested process domains in Earth history, and as much more recent research priorities among environmental scientists. The United States National Research Council's (USNRC) (2001) report defined the Critical Zone as the "heterogeneous, near-surface environment in which complex interactions involving rock, soil, water, air and living organisms regulate the natural habitat and determine availability of life-sustaining resources." This is commonly identified as "the fragile skin of the planet defined from the outer extent of vegetation down to the lower limits of groundwater" (Brantley et al., 2007, p. 307). Shortly following the USNRC report, a collaborating body of Earth scientists initiated what was then called the Weathering Systems Science Consortium (WSSC) (Anderson et al., 2004), intent on studying "Earth's weathering engine" in the context of the Critical Zone. The WSSC evolved into a more-encompassing Critical Zone Exploration Network (CZEN) as the collaboration involved more ecologists, hydrologists, and pedologists less intent on examining the weathering engine. Still CZEN has been and is a significant locus of weathering research. Not to be ignored, however, is the longstanding pursuit of regolith science ( Scott and Pain, 2008 ; Taylor and Eggleton, 2001 ; Cremeens et al., 1994 ; Ollier and Pain, 1996 ), more or less parallel to that of the Critical Zone although not named as such. Ollier and Pain (1996 , p. vii) outlined in the preface of their book goals for "regolith studies" that are nearly exactly those espoused by CZEN for the same environment. In the effort to draw together disparate research approaches, this chapter on weathering systems will draw on a variety of sources and scientists who study weathering in the Critical Zone/regolith: geochemists, geomorphologists, soil scientists, hydrologists, petrologists, mineralogists, and others. Viles' (2013) discussion of weathering system synergies elucidates the environmental as well as disciplinary entanglements, pointing

out disparities in questions asked and methodologies used, despite the convergence on a congruent topic of investigation.

Yatsu's (1988, p. 2) definition for "weathering" remains the simplest and most flexible: "Weathering is the alteration of rock or minerals in situ, at or near the surface of the earth and under the conditions which prevail there." Indeed, rocks and minerals do thermodynamically adjust within the active Critical Zone envelope. Yet, the term "weathering" implies a predominance of atmospheric control over processes, when in actuality the inherent properties of the rock and mineral may be foremost (Hall et al., 2012). Neither is the atmosphere itself the sole nor necessarily the dominant weathering agent. Hall et al. (2012) forwarded a replacement term, "rock decay." The same term was used by early geomorphologists (Merrill, 1906 ; Chalmers, 1898 ), and Ollier (1963) referred to chemical weathering as one of several types of "rock decay." Hutton (1788), in his famous paper, never used the term, "weathering," instead used "decay" (although the cause of which he generally ascribed to "sun and atmosphere"). The term, "decay," appears commonly in the stone conservation literature (e.g., Prikryl and Smith, 2007). A growing body of "weathering" scientists is following this example (cf. Dorn et al., 2013a).

Even so, "rock decay" is not quite sufficient to cover the realm of weathering of all earth surface materials. Materials need not be rock to decay. Unconsolidated sediments can decay in situ, in temporary surface storage prior to burial and cementation. Direct byproducts of surface weathering may also weather. For instance, neoformed duricrusts precipitated in the soil (Nash and McLaren, 2007 ) may undergo decay when environmental conditions modify, in which the precipitate is no longer stable in its present form. For illustration, Eren and Hatipoglu-Bagcit (2010) reported karst-like solution features in exposed hardpans of caliche in Turkey. Calcium carbonate precipitated in the soil in dry conditions, but the now-exhumed hardpan was exposed to solution from rainwater. Twidale (1987) explained karst-like sinkholes in laterized quartz and clay in tropical Australia that resulted from Holocene drying as well as biogenic decay agents.

Regolith is the "unconsolidated or secondarily cemented cover that overlies more coherent bedrock... formed by weathering, erosion, transport and/or deposition of older material. Everything from fresh rock to fresh air" (Scott and Pain, 2008 , p. 427). Regolith therefore coincides with all of the Critical Zone apart from the subaerial vegetation, and happens to be the natural focus of rock and mineral decay. If the regolith develops from in situ bedrock, the term saprolite is sometimes used for decayed bedrock that retains the original rock structure and fabric. Saprock is the slightly weathered transitional form between saprolite and rock, whereas some authors use the term saprolith to cover the combination of saprolite and saprock (Scott and Pain, 2008). Finally, soil is the uppermost segment of the regolith generally (but not always) with incorporated decomposed biotic components as well as genetic horizon organization. Soil is at the direct interface of Earth materials at the atmosphere, surface water, and biosphere, and covered in this text by Dixon (Chapter 5). Figure 4.1 diagrams the profile of the regolith, otherwise known as the terrestrial portion of the Critical Zone, whereas Fig. 4.2 illustrates a real-world example, complete with its vegetative cover.

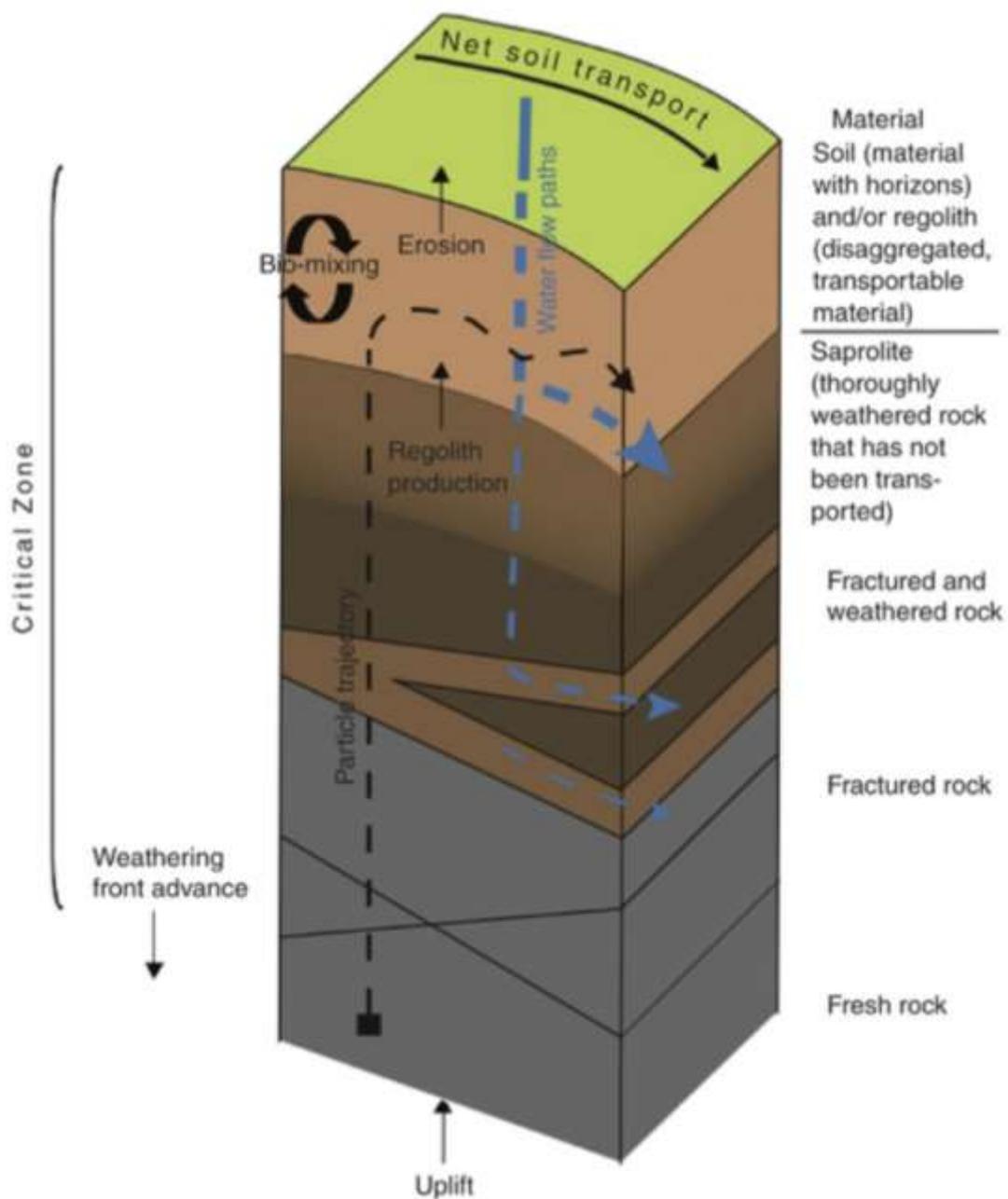


Figure 4.1 Profile of the terrestrial portion of the Critical Zone, or regolith, as conceptualized by Anderson et al. (2007). Penetration of the weathering front progresses from surface downward. Rock particles liberated by weathering have a net movement upward, then at the surface mixed and moved laterally and downslope.



Figure 4.2 A regolith profile, in decayed schist and gneiss, in a road cut near Evergreen, Colorado, USA. Relative scale is indicated by the nearby trees. Less resistant rock accounts for the concavities in the profile, while more resistant corestones and quartz veins stand out. Iron oxide staining is seen primarily along fracture zones and around corestones. The irregularity of the decay profile is evident, with highly decayed portions against less decayed portions incongruently, as identified by Phillips (2001b) .

## **4.2. Weathering relevance to other Critical Zone processes**

Although a number of biological, atmospheric, and hydrologic processes operate and interact in the Critical Zone, the interfaces with the lithosphere illustrate the importance of weathering in the pedosphere and regolith, connected within the Critical Zone. The breakdown of rocks and minerals and the mobilization of elements are relevant to pedogenesis, water quality, atmosphere and climate, mineral resources, and life itself (Buss et al., 2013).

### **4.2.1. Pedogenesis**

Soil is the literal foundation for terrestrial life, providing a physical platform, a water–air–nutrient reservoir, and a primary means to recycle both inorganic and organic elements for future use. Decay of organic material is beyond the focus of this chapter, although it is important to consider that almost all soil types have an organic component, and some byproducts of organic decay become rock/mineral decay agents. The processes of pedogenesis are well covered in Birkeland (1999), Schaetzl and Anderson (2005), and Phillips (1993, 2001a ), and couched in the factorial controls on pedogenesis first expressed by Jenny (1941) and Simonson (1959) . For the breakdown of rock and mineral, soil forming factors are essentially identical to those of weathering (Pope et al., 1995): the nature of the parent materials, influx of new material and efflux of byproducts, the atmospheric conditions over the timespan of the process (including temperature and moisture availability), the presence and type of biotic activity, the availability of abiotic and biotic weathering agents, and time. Pedogenesis in the Critical Zone is discussed in greater detail by Dixon (Chapter 5).

### **4.2.2. Water Quality**

The transfer from atmospheric water to ground- and soil-water, and surface water takes place within the soil and regolith. Although diagenesis and hydrothermal processes below the Critical Zone impart dissolved solutes to groundwater, Critical Zone water also plays a role. With the assistance of water, rock-decay processes add dissolved solutes and suspended particulates deriving originally from rock. The net flow of water and solutes in the Critical Zone depends on the hydraulic gradient, downward under the influence of gravity (flushing components into the groundwater), or upward with strong vegetation uptake or evaporation (in which precipitates such as calcium and salts become part of the soil).

A segment of Critical Zone studies involves rock-decay processes and surface- and groundwaters interacting with polluted environments, such as exposed quarries, mine spoil, urban landfill, and agricultural soils. For instance, the problem of acid mine drainage involves water in contact with iron sulfide minerals in newly exposed rock (as mine spoil or as bare rock). One byproduct is sulfuric acid, itself a potent weathering agent, and a detrimental additive to aquatic life unaccustomed to low pH (Raymond and Oh, 2009; Bond et al., 2000). In the example of urban and industrial contaminants, biogeochemical surface reactions, essentially weathering processes although not necessarily identified as such in this body of research, are a key point in the analysis and remediation of polluted soils (cf. Trindade et al., 2005; Urum et al., 2004).

### **4.2.3. Supporting Life, Conditioning Ecosystems**

Although ongoing research continues to identify a growing number of extremophiles (Rothschild and Mancinelli, 2001), the “normal” terrestrial ecosystem

relies on a Critical Zone system that regulates and supports life, including recycling dead organisms, storing and making available nutrients and water, and physically supporting and sheltering organisms. Weathering processes help to provide nutrient elements out of decomposed minerals, and to produce soils that are acceptable to organism regeneration. The soil itself, as a product of weathering, strikes a balance of being porous enough to allow transfers of water, air, and nutrients, but also substantial enough to literally support the weight of vegetation or remain intact when tunneled by burrowing animals.

Members of the ecosystem present a feedback of weathering to the Critical Zone. Biotic agents contribute a significant, if not dominant, share of weathering work to the Critical Zone, involving multiple factors. First, root respiration is responsible for high concentrations of CO<sub>2</sub> in soil air, which in turn combines with water to form carbonic acid (Brook et al., 1983). Second, byproducts of organic decay directly contribute organic acids and chelates, discussed later. Third, penetration of roots and burrowing animals provide pathways for weathering agents and increased exposed surface area. Fourth, among the least understood, are the many symbiotic microorganismal, chemical reactions that exist in the pedosphere (cf. Leake et al., 2008; Chorover et al., 2007). Finally, there are newly discovered impacts of organisms beyond the classifications enumerated earlier. One such phenomenon is the verified enhancement of dissolution of Ca and Mg silicates associated with ant colonies (Dorn, 2014). Through mechanisms not yet understood, although perhaps involving geophagy of olivine and plagioclase, the process combines CO<sub>2</sub> to form pedospheric magnesium- and calcium carbonates, while acting as a terrestrial carbon sink.

#### **4.2.4. Regulating the Environment and Climate**

One of the most significant hypotheses put forth in weathering system science has been the role of weathering and the sequestration of carbon, and thus regulation of greenhouse gasses and climate. Originally developed by Raymo (1989), following a supposition by Chamberlin (1899), the hypothesis states that weathering of fresh silicate rocks (high on Reiche's (1950) weathering potential, see Figure 4.3) will draw down atmospheric carbon dioxide. Further, episodes of increased orogeny on a planetary scale (such as the Alpine orogeny) expose more fresh silicate rock, thereby allowing more weathering and more CO<sub>2</sub> drawdown, enough to precipitate global cooling. Indeed, the onset of large-scale orogeny does coincide with global cooling over late Cenozoic time, as well as earlier periods of coupled orogeny and climate change (Raymo and Ruddiman, 1992; Raymo, 1989). Originally beset by difficulties dating and quantifying the degree of weathering, the hypothesis has better support in recent work, by Raymo and her colleagues (cf. Raymo et al., 1997) as well as independent researchers (cf. Dupré et al., 2003; Kump et al., 2000; Liu et al., 2011). Still, some researchers find conflicting evidence. Dixon et al. (2012), working in the xeric climate of California, found evidence to suggest the opposite, tectonic/weathering forcing that increased orogeny beyond a certain threshold may retard chemical weathering. Nevertheless, the premise is strong justification for ongoing research, with age control, quantified rates and degrees of weathering, but over different climatic and tectonic settings.

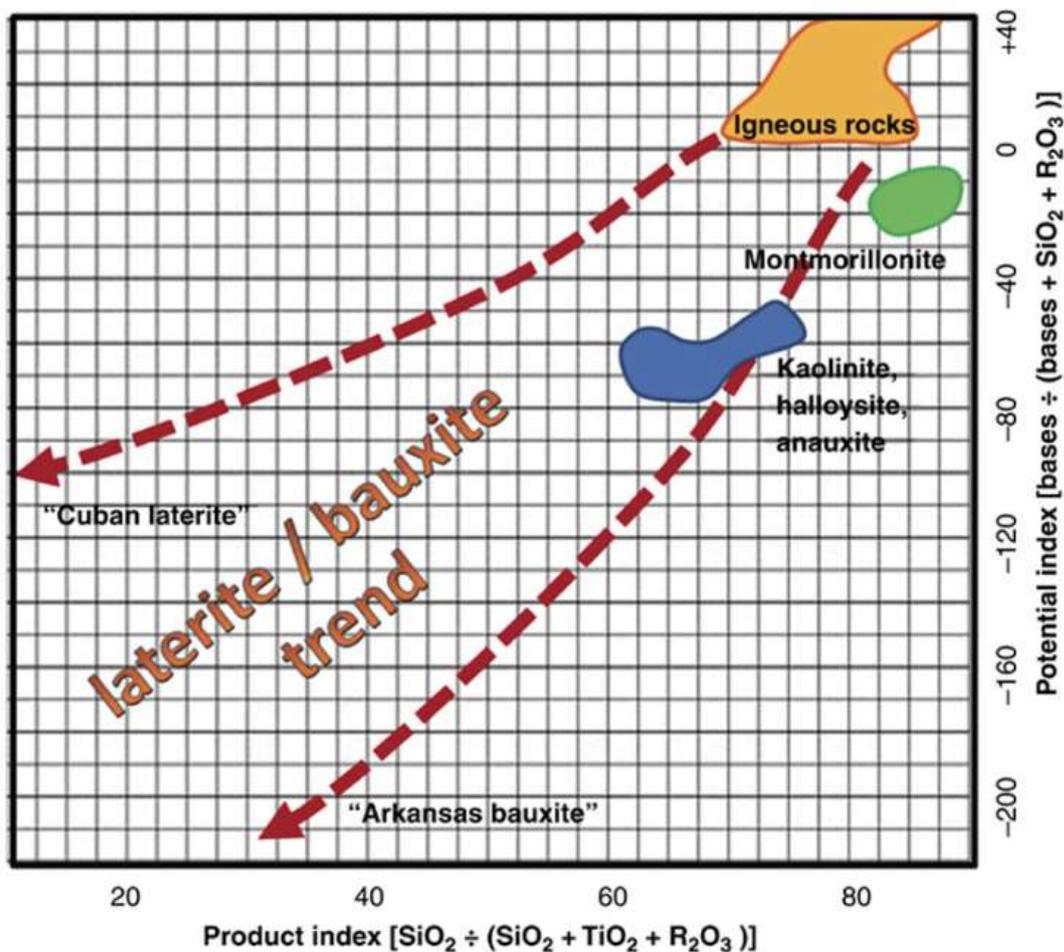


Figure 4.3 Reiche's (1950) weathering product versus weathering potential indexes, showing the trend from igneous rocks to end-stage bauxite and laterite.

#### 4.2.5. Providing Natural Mineral Resources

The byproducts of weathering are occasionally useful as natural mineral resources, and their surface presence enhances their economic availability when compared to deeper strip mining, quarrying, or tunnel mining. Materials such as gravel and aggregates, glass sand, clay, and aluminum, magnesium, and iron ores may occur at the surface as end-products of weathering (Ollier and Pain, 1996; Taylor and Eggleton, 2001). Likewise, precious and strategic metals such as nickel, gold, silver, and copper tend to concentrate in the upper reaches of well-weathered regolith, or may occur in placer deposits downstream of regolith source areas (Ollier and Pain, 1996; Taylor and Eggleton, 2001). Australia has been the focus of elaborate exploration of regolith resources, including geophysical methods and remote sensing to identify regolith structure as well as chemistry (Butt et al., 2008; Pain, 2008). Wilford (2011a) employed airborne gamma-ray spectrometry to assess the presence of potassium, thorium, and uranium, as part of a weathering intensity index, resulting in a continent-scale mapping of weathering intensity for Australia.

Soil is of course a vital natural resource and byproduct of weathering. The roots of the science of pedology are found in initial attempts to understand soil and manage it. Soil is subject to both natural and human impacts. Soil management and impacts are beyond the scope of this chapter, but Richter (2007) provided an overview of the history and future human interfaces with soil resources.

### **4.3. Types of weathering (rock decay)**

The traditional view of rock weathering in textbooks separates mechanical and chemical weathering processes. The former involves breakdown of the rock or mineral into smaller pieces without changing its chemical composition, whereas the latter involves breakdown of the rock or mineral by means of changes to the chemical compositions (adding or subtracting atoms, ions, or molecules, at the smallest scale). Many authors consider biological weathering to be a separate third type of weathering. Indeed, biological agents are important and in many cases dominant, although ultimately chemical or mechanical in nature. At the nano-scale of the mineral surface (Dorn et al., 2013b), the creation of smaller particles – by mechanical or chemical means – is almost semantic. Depetris et al. (2014) state that mechanical weathering is first to occur as the opening or exposure of surface area to subsequent chemical processes. Opening surface area to chemical alteration is accurate, although it may also be argued that the mechanical flaws in the rock or mineral that represents a locus of failure may well have been initiated by chemical attack. In reality, chemical and mechanical processes of decay, including the biological types, are synergistic, occurring in parallel and often with positive feedback between processes. Blair (1975) summarizes the relationship well: “It is frequently difficult and unnecessary to separate physical weathering from chemical weathering, for the two usually work together.” A summary of the processes is provided here.

#### **4.3.1. Normal Stress**

Stress exists where pressure or force exceeds the ductile capacity of the rock or mineral. Normal stresses include compressive and tensile forces. The simplest example of compressive stress is that of the mass of a heavier object resting on a rock, which may induce fracture. Tectonic forces of compression also impart fractures, even at the crystal level; these fractures are in turn avenues for weathering by other agents.

A common subject of study in the rock decay/weathering literature is the pressure exerted by substances expanding within confined spaces in the rock or mineral, such as in cracks or voids. Ice is a common example, where the volume of frozen water increases over its liquid state. In confined cracks, pressure exerted by the ice widens the crack; the process may be repeated as more liquid water is allowed in and also freezes. It is argued (Bland and Rolls, 1998), however, that confining pressures may not be as much as anticipated. First, ice is capable of compressing in a ductile fashion. Second, cracks are almost always actually planes of weakness within the rock, capable of spreading the stress laterally instead of against the walls of the crack. Water pressure in advance of ice crystal growth can be more important (Bland and Rolls, 1998). Alternately, thermal stress (given later) may in fact account for mechanical fracture observed in cold (Hall and Hall, 1991) and hot (Viles, 2005) condition.

Salt-crystal growth is similar in principle to ice-crystal growth, although with less controversy. Salts are common in dry regions but also present near ocean coasts, in urban areas, and as long-transport aerosols (Goudie and Viles, 1997; McDowell et al., 1990), so more prevalent than would be expected. A substantial body of research on salt-induced decay derives from stone conservation and cultural heritage studies. Salts enter the rock in water solution, and then precipitate out as growing crystals as the water evaporates, so some duration of drying is required. Repeated wetting and drying is capable of inducing cycles of salt-crystal stress on a rock, more capable of

damage to the rock. Not only at the scale of visible fractures or crystal/detrital grain boundaries, salt-crystal stress is also observed at the micro- and nano-scale, rupturing mineral weaknesses and lattices (Pope et al., 1995). Calcite crystal growth from calcium carbonate deposition in dry areas would present similar mechanical stress in soft or decayed rock (Boettinger and Southard, 1990).

The growth of plant roots has long been associated with mechanical weathering (cf. Ollier and Pain, 1996; Yatsu, 1988), and recent studies attribute the importance of both large and fine roots to mechanical weathering in the soil (Gabet and Mudd, 2010; Richter et al., 2007). Many introductory textbooks feature illustrations of boulders pried apart by tree roots, an image that would seem obvious to accept. Bland and Rolls (1998, p. 159) argued that the pressure of plant roots is insufficient to fracture rocks except in very weak rock, given that the radial pressure is only exerted onto two rock planes in a crack. This posits that plant roots simply take advantage of existing fractures initiated by other mechanical processes; the appearance of physical stress is simply coincidental. Root stress may be one of many processes where field-observed phenomena have yet to be rationalized by laboratory or numerical modeling. Plant roots do play a role in disturbing the soil, including wind-thrown root masses that may rip up weakened regolith rock (Gabet and Mudd, 2010; Phillips and Marion, 2006). Researchers appear to agree that microscale root hyphae are capable of mechanical breakup of minerals, attributed to lichens, algae, and fine roots of vascular plants (Richter et al., 2007; Duane, 2006; Chen et al., 2000; Lee and Parsons, 1999; Hall and Otte, 1990). Mechanical processes include both compression stress by root hydration as well as the tensile “ripping up” of attached particles by roots or hyphae.

Degrees of mineral expansion and contraction are possible with wetting and drying. Mineral hydration can expand the crystal dimensions, exerting stress similar to that of salt. The process of hydration is a chemical one, in which hydroxide molecules are incorporated into the mineral, thereby changing its chemical composition. However, the process also resembles a mechanical stress at the molecular scale. Hydroxide ions disrupt the mineral lattice, rendering it weaker in some cases, as in silica hydration (cf. Aomine and Wada, 1962), or expanding and splitting the crystals, as in salt, clay, or mica hydration (Velde and Meunier, 2008; Doehne, 2002). Hydration can take place in any environment where at least hygroscopic water is present. Dehydration works in the opposite fashion, removing hydroxide ions from hydrated minerals, which has the potential to collapse the mineral structure.

Salts, iron minerals, calcite, clays, and even amorphous glass (e.g., in volcanic ash) can undergo hydration and expansion. The reported mechanical stress of “fissuresols” (Villa et al., 1995) probably derives from clay-mineral expansion and calcite crystal growth (Dorn, 2011). Clay hydration is a good example of a chemo-mechanical process: Water bonds to the phyllosilicate crystal lattice, this in turn causes the lattice to expand, exerting physical pressure.

Apart from the above-mentioned root-ripping, tensile stresses are relatively less relevant to rock decay. One important exception is the process of pressure unloading, also sometimes known as sheeting, dilation, and exfoliation. 1 Blackwelder (1925) is credited with the first comprehensive description of pressure unloading. Where large rock masses, formerly buried under tonnes of crust, regolith, or glacial ice, are exposed to the atmosphere, confining pressures are relaxed, and the rock expands. Where it exceeds the plastic expansion limits, the rock fractures parallel to the exposed surface. Perpendicular fractures also develop because rock is

incapable of stretching to its new expanded volume. Domed rock structures as well as cliff faces experience unloading stress. Prior to separation and falling, the widening fractures can channel water into the rock, along with biotic and other external materials, contributing to further weathering. The released rock slabs fall and become components of the detrital geomorphic system (Pope, 2013a).

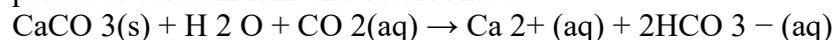
#### **4.3.2. Thermal Shock and Fatigue**

The role of thermal shock, of rapid temperature excursions within the rock, has been a longstanding debate in the study of mechanical rock decay. Early researchers such as Blackwelder (1933) and Reiche (1950) dismissed the importance of thermal shock in producing rock fragments. However, with field data indicating the extremes of temperature change that exist on rock faces, Hall and Hall (1991) demonstrated the potential for tensile contraction forces during temperature excursions in dry polar rock surfaces. Viles (2005) presented comparable data from a hot desert situation. In fact, Sumner et al. (2004) found similar ranges of temperature excursions in deserts of South Africa and Antarctica, capable of thermal shock weathering. Repeated thermal excursions could result in eventual material fatigue (Hall, 1999).

A subset of thermal shock could include explosive detachment, caused by rapid expansion due to extreme heat. Numerous researchers note spalling and granular separation after fires (Dorn, 2003; Ollier, 1983a; Blackwelder, 1927), and the explosive expansion of water vapor just below the rock surface is a likely cause. The prevalence of natural fires in a number of ecosystems (Belcher, 2013) suggests that fire weathering may be more important than previously suspected. Even more extreme than fire, Knight and Grab (2014) make a good argument for the occurrence of lightning weathering, certainly qualifying as explosive, on exposed mountain ridges. They suggest that in some locations, lightning-induced weathering is a better explanation for shattered rock previously attributed to ice weathering processes.

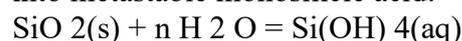
#### **4.3.3. Solution and Dissolution**

Solution and dissolution involve the ability of water, often fortified with acids, to remove ions. Solution is the most straightforward process, able to disrupt molecules in one step by creating a solute ion and an aqueous byproduct. Solution of calcium carbonate in the presence of acid precipitation or acidic groundwater produces calcium ions and bicarbonate:



This process is relevant over a broad portion of the planet where carbonate rocks are present, and responsible in the extreme case for karst landscapes, but also for the common calcium load in surface waters.

Solution of silica,  $\text{SiO}_2$ , is also possible, although at much slower rates than solution of carbonate rocks. In the presence of water, silicate molecules dissociate into metastable monosilicic acid.



Solution is the process by which quartz, among the most resistant of minerals, is capable of decay. In its most extreme form, silica solution is responsible for silica karst observed in sandstone and quartzite rocks (Wray, 2013). Such landscapes have been otherwise stable enough, in relative tectonic quiescence, to be exposed to the long-term effects of solution. At much smaller scales, silica solution is probably responsible for overgrowths and glazing observed on quartz sand particles (Krinley and Doornkamp, 2011).

Dissolution is a more complex process, but important to the silicate minerals other than quartz. Acid is involved, such as carbonic acid by atmospheric precipitation or by combination of water with soil, air, organic acids near the soil, and around sulfide deposits (including mine waste). For example, the feldspar mineral albite decays in weak acid to the aluminosilicate clay, kaolinite, releasing also quartz (often as a precipitated silicate cement or amorphous coating) and sodium ions in solution:



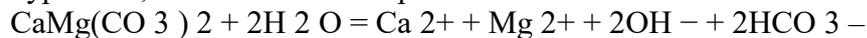
Hydrolysis, sometimes also known as incongruent dissolution due to its multistep and parallel decomposition sequence, is also relevant to the decomposition of silicate minerals. It may take place with acids or with water. For aluminosilicate minerals, the process results in end-product clays with ions of potassium, sodium, calcium, and quartz precipitate (dissociating from an initial silicic acid byproduct), as in this example for an alkali feldspar:



Not all silicate decomposition results in the formation of clay, as in this example of olivine, but does produce iron and magnesium ions (which may precipitate out), as well as precipitated quartz (again dissociating from a silicic acid byproduct):

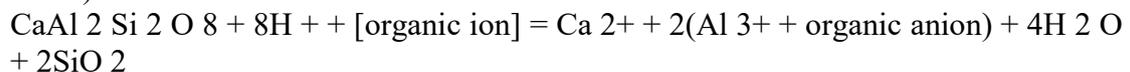


Dolomite decomposition is an example of hydrolysis outside of the silicate minerals, producing calcium and magnesium ions as well as carbonic acid as a byproduct, which in turn take part in further solution reactions:



#### 4.3.4. Rock and Mineral Decay by Organic Molecules

Organic chemicals that derive from living as well as decomposing dead organisms introduce unique and sometimes particularly intense decay agents into the near-surface regolith. The soil, particularly in the rhizosphere of root-soil interaction, is highly active chemically, including oxidation, solution, hydrolysis, and chelation reactions. Ollier and Pain (1996) pointed out the role of oxidation in plant uptake of iron and other nutrients, and the role of bacteria in silica depletion. Organic acids such as oxalic acid and humic acid can be effective in the decay of silicate minerals, such as this example of anorthite decomposition. (Scott and Pain, 2008, after Trencases, 1992):



The reaction releases Al as an organic compound, as well as calcium ions and a quartz precipitate.

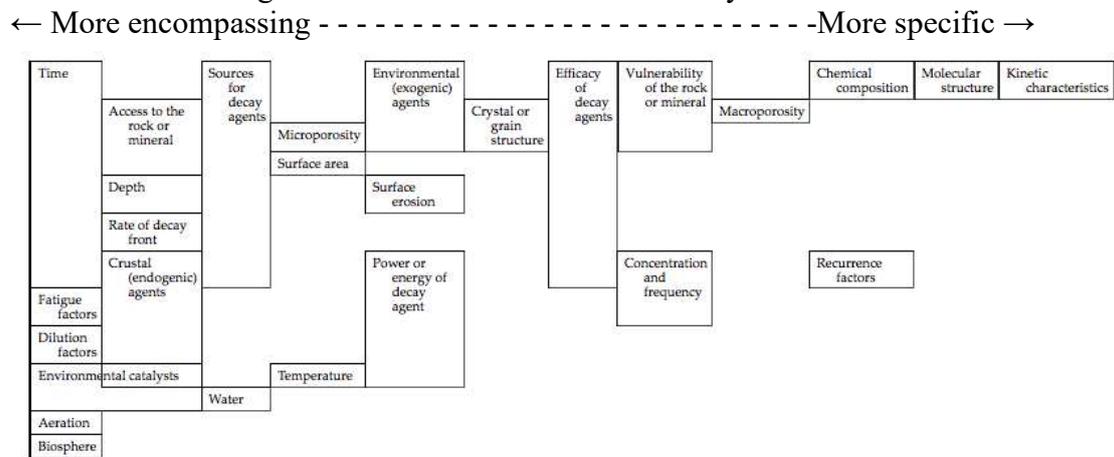
Chelation is the process by which metals such as aluminum and iron are preferentially extracted by acids and organic molecules (such as ligands) derived from decomposing vegetation. Lichens, as pioneer organisms at the lithosphere/biosphere interface, incorporate chelation and oxalic acid in initial rock decay (Wilson and Jones, 1983; Schatz, 1963). Masiello et al. (2004) attributed chelation processes for the translocation of iron and aluminum with carbon in grassland soils.

#### 4.4. Factors relevant to rock decay

The degree and rate of rock decay is influenced by three arch-factors: (1) time, (2) availability of rock and mineral decay agents, and (3) efficacy of rock and mineral

decay agents, indicated in Table 4.1. Rock decay tends to be a slow process, but decay increases with time up to the point of weathering-agent saturation or end-product stability. With enough time, extensive rock decay can take place, anywhere. Time is therefore the determinant factor for all others. The weathering system also requires available decay agents, and materials to decay (assumed, not indicated in Table 4.1). Exogenic decay agents derive from the environment, whereas endogenic decay agents derive from the crust itself (such as gravity, tectonic stress, geothermal heat). Decay agents are in turn rendered more or less effective by intrinsic properties of the rock or mineral, and by factors in the environment that catalyze or diminish the power of the decay agent. Additional sub-factors become relevant specific to the time and location. The environment is clearly important, but equally so is the nature of the rock and mineral.

Table 4.1 Cascading Factors of Rock and Mineral Decay



Derived from multiple sources, including Viles (2013, 2001), Anderson et al. (2007), Phillips (2005), Pope et al. (1995), Simonson (1959), and Jenny (1941). The factors listed are encompassing and representative, although not necessarily exhaustive, as additional factors may be found.

Reiche (1950) introduced a concept of “weathering potential,” useful to distinguish different materials with respect to eventual thermodynamic equilibrium, therefore a chemical weathering attribute. Materials with a high weathering potential consist of freshly exposed igneous rock, most out of balance with the ambient surface conditions of the Critical Zone. Materials with low weathering potential are near the end-stage of weathering, close to thermodynamically stable at the Critical Zone, such as bauxite and laterite. Reiche illustrated the weathering potential index (leachable bases compared to silica) graphed against a weathering product index (silica loss compared to more stable titanium and aluminum oxide), reproduced in Fig. 4.3 .

It should not be surprising that different factors dominate in different settings; no single factor is capable of distinguishing or determining the degree of decay. This applies spatially, across regions and landscapes or at smaller scales such as slope elements, and vertically through the regolith profile. Climatic factors provide one example. Studies by Dixon et al. (2009) and Brady et al. (1999) show the importance of climatic variables such as rainfall, especially when combined with biotic weathering agents, in sampling transects over local to regional scales. However, Pope (2013b) , Pope et al. (1995) , Viles (2013) , and Anderson et al. (2007) provide evidence for factors such as soil and regolith structure, overburden erosion rates, and differences in weathering agents over-ride presumed dominance of enhanced decay due to high temperatures and water availability in the tropics. Likewise, Dixon and Thorn (2005) reported the dominance of chemical processes in subarctic and alpine environments ordinarily presumed to have greatly dampened chemical activity. Decay

agents and factors vary with depth as well. Anderson et al. (2007) illustrated the terrestrial Critical Zone profile (the regolith, reproduced in Fig. 4.1) with chemical or mechanical processes dominant nearest bedrock, and chemical processes dominant in the upper saprolite, whereas a mix of mechanical (including transport), chemical, and biological processes impact the soil. Anderson et al. pointed out that this is a generalization, where variation in these trends may be the result of weathering potential and weathering agent efficacy. Further, it is important to remember that weathering processes are, as mentioned earlier, synergistic, and not strictly only chemical or only mechanical. At the depths of the Critical Zone, the transition to the diagenesis-influenced crust, mechanical stresses help to break up the massive rock formed in the crust, yet chemical processes are not absent and may still take advantage of existing petrographic weaknesses, which may in turn fail under physical stress.

Scale is relevant to rock decay factors. Rock decay initiates with mineral decay, the smallest of building blocks within the rock (whether they be interlocking crystals or cemented detrital particles). Mineral decay ultimately rests on the nanoscale separation of molecules, either by chemical or mechanical means. Factors influencing the processes at the nanoscale “boundary layer” (Dorn et al., 2013b) may be quite different from those of ambient atmospheric conditions at the ground surface (Pope et al., 1995). Present “weather” and “climate” of the location may well be poor proxies for the actual nanoscale boundary layer. That said, the nanoscale boundary layer is not divorced from its environmental envelope. Nested environmental realms exert influence (Table 4.1), if not control, on successively smaller envelopes, from planetary and regional to local to site-specific and finally to the mineral surface environments.

Figure 4.4 a and b illustrates the weathering/mineral system, including global factors but focusing on the weathering boundary layer. Note that both sides of the system, the mechanical and the chemical, are linked through a feedback “tunnel” indicated in each diagram. The system as illustrated is conceptual in energy contributed and used for weathering, borrowed in part from the ecosystem energetics of Odum (2000) and the factors illustrated in White et al. (1992). The combined lithosphere, hydrosphere, atmosphere, and biosphere (the terrestrial inputs) provide the chemical weathering agents, whereas the intensity and rates of these agents are modified by both external energy contributions and parent material properties (including mechanical processes). Likewise, mechanical weathering agents are sourced externally, and modified by external (such as geomorphic) and internal (rock/mineral) properties, as well as chemical weathering. As the process evolves, energy feedback (both positive and negative) moves into the system, as well as some net energy loss as the system progresses toward stability.

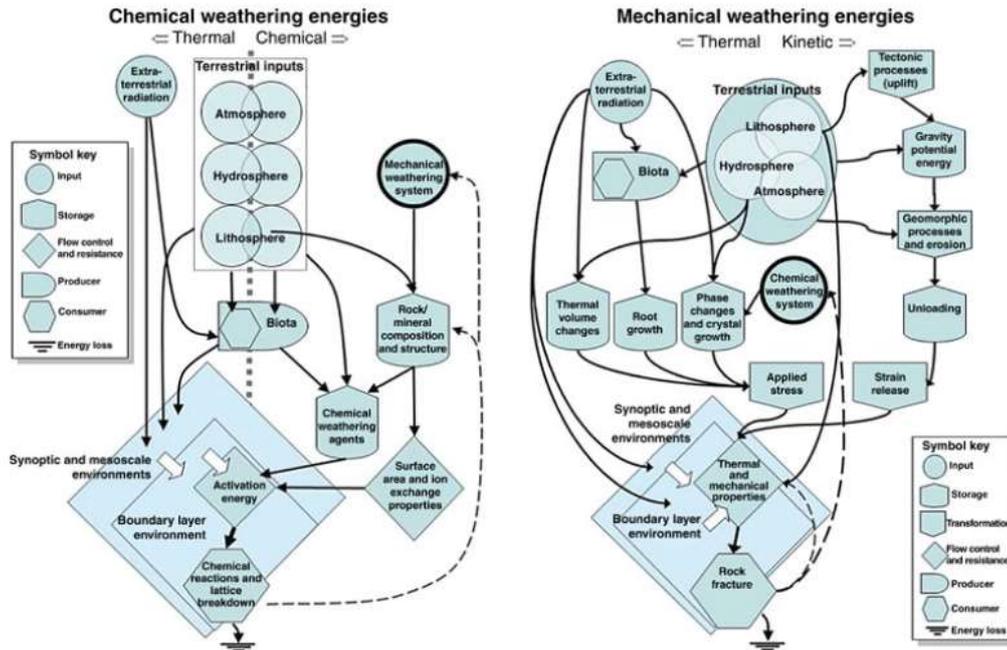


Figure 4.4 The weathering (rock decay) system, expressed in terms of cascading energy. Chemical (2a) and mechanical (2b) processes and factors are treated separately, but are connected by a feedback tunnel indicated in each diagram. Symbols follow the standard typology of systems.

The concluding message is that it is possible to conceptually model decay agents and factors, but most researchers on the Critical Zone recognize that individual cases and locations are quite varied and often surprising. This is a reality that challenges our abilities to model the weathering system and to reconcile models with field observations (Dorn et al., 2013a).

#### 4.5. Rock decay in three dimensions: the “weathering mantle”

Rock decay is readily apparent at the surface, with direct exposure to the elements, and at the near-surface, associated with pedogenesis. The Critical Zone is often much deeper (Fig. 4.1), and it is argued (Buss et al., 2013, for instance) that significant chemical decay is capable 10 m to more than 100 m below the surface. Terminology used in the weathering mantle varies; Table 4.2 illustrates the regolith zones and terminology used by several authors, suggesting congruence but some overlap. It should be noted that the terminology and classification is biased toward crystalline bedrock, and in humid-warm/hot environmental conditions (presently or presumed in the past). Figure 4.2 illustrates a typical regolith profile, in this case in weathered metamorphic rock.

**Table 4.2**  
**Segments of the Regolith or Weathering Mantle as Described by Different Authors**

Authors	Pavich et al. (1989)	Ackworth (1987)	Jones (1985)	Eggleton (2001) and Migoñ (2013)	Anderson et al. (2007), see Fig. 4.1
Regions	Appalachian Piedmont	Humid tropics	Humid tropics	"Idealized regolith"	Vertical Critical Zone
Increasing depth	Soil	A horizon	Quartz-rich soil	Pedolith	Soil and/or regolith
	Massive subsoil	B oxic horizon	Duricrust		
	Saprolite	Upper saprolite: sandy-clay/clayey-sand	Completely decomposed rock	Saprolite	Saprolite
	Mid-saprolite: massive secondary clays		Decomposed rock, original structure		
	Lower saprolite: altered rock	Weathered rock, some pseudomorphs	Saprock		
Weathered rock	Transition bedrock, fractured		Slightly weathered rock	Fractured + weathered rock	
Fractured rock					
		Weathering front	Weathering front	Weathering front	
Fresh rock	Fresh rock		Bedrock	Fresh rock	

In the deepest regolith, the influences of the surface (meteoric water, ambient surface temperatures) transition to those of the crust (geothermal heat, diagenetic processes), although in most cases, the Critical Zone derives minimal direct deep earth influence. Temperature ranges reach near-equilibrium at depths of 1 m or less (Anderson, 1998). The boundary between diagenetic and weathering processes is ill defined, and perhaps semantic. Ollier and Pain (1996, p. 72) rightly point out that decay occurs in both water-saturated (e.g., below the water table) and water-unsaturated zones; this entire thickness is encompassed in the defined "Critical Zone." Precipitated elements that would derive from hydrothermal interactions help distinguish hydrothermal alteration from surface-derived weathering (cf. Thuro and Scholz, 2003 ; May, 1994 ). Even so, the results are not straightforward, and sometimes controversial (Ollier, 1983b; Young and Dixon, 1983).

Deep regolith profiles are noted in tropical and subtropical locations, indeed one of the Critical Zone Observatories is established in the deeply weathered terrain of Luquillo, Puerto Rico, for the purposes of observing deep decay processes. Although conditions favoring intense and effective decay are present in tropical environments – abundant water, organic compounds, and high temperatures – time and stability are also relevant to developing deep profiles. Deep regolith profiles do occur outside the tropics, and although some argue that these can be relict of past tropical times over the long durations of regolith formation (by climate change or tectonic drift), deep weathering need not be tropical in origin (Pope, 2013b). Blair (1975) , for instance, described deep grus in some places more than 50 m thick in Pikes Peak granite of the Colorado (USA) Front Range. In prestressed rock fabric, groundwater was responsible for chemical deterioration of biotite, lending to expansion and disaggregation. Although Chapin and Kelley (1997) presumed deep chemical weathering in the Pikes Peak granite to a warmer, wetter climate on the Eocene-aged erosion surface, Blair (1975) found the profile devoid of clay mineral formation characteristic of feldspar decay seen in the tropics. Bazilevskaya et al. (2013) summarized 27 studies in both temperate and tropical climates and found that although granitic terrain regoliths of similar age were slightly thicker, they were not

statistically so, and regoliths of similar age on rocks of basaltic composition were thinner in tropical climates. (These results parallel the relative lack of variation in chemical denudation rates across climatic regions, noted by Saunders and Young (1983), and summarized and discussed in Pope et al. (1995).) Dosseto et al. (2011) showed that soil production rates (regolith in general) in granitic parent materials are “relatively insensitive to climate.”

Regolith thickness can be thought of as a balance between regolith production by weathering at the saprolite/rock interface (the weathering front), and loss of material by erosional stripping at the surface. Ahnert (1987) modeled the production of regolith thickness determined by weathering rate and differentiation between chemical and mechanical weathering efficacy ( Fig. 4.5 ), and precedes a similar conclusion by Gabet and Mudd (2009) by more than 20 years. In effect, regolith thickness influences the rate of weathering. The model predicts a “critical thickness” ( $C_c$ ) at which rates of weathering peak, when there is at least some chemical weathering in the system (e.g., almost always). In systems of purely mechanical weathering (rare), weathering rates decrease continuously as disaggregation relieves the stress on solid rock, and further mechanical weathering is unlikely to have continuing impact. For systems with combined mechanical and chemical weathering (almost all) or only chemical weathering (also rare), weathering rates increase from initial exposure up to the critical thickness point, owing to increasing porosity in the regolith and ability to maintain water. In a previous study, Ahnert (1976) described a zone of “optimal chemical weathering” somewhat below the surface which decreased as overburden cover increased (also observed in recent studies, such as Brantley and Lebedeva, 2011, and Bazilevskaya et al., 2013). The modeled relationship (see also Fig. 4.5) is as follows, for degree of weathering per unit time,  $W$  (an overall weathering rate), from the onset of weathering until the critical thickness has reached ( $C < C_0$ ):

$$\text{for } C < C_c, \quad W = W_c \left( 1.0 + k \frac{C}{C_c} - \frac{C^2}{C_c^2} \right)$$

where  $W_c$  is the chemical weathering rate,  $k$  is a coefficient,  $C$  is the regolith thickness, and  $C_c$  is the specific critical regolith thickness. Smaller values of regolith thickness  $C$  compared to the critical thickness  $C_c$  equate to lower unit weathering. Beyond (thicker than) the critical thickness ( $C > C_0$ ), weathering rates decrease with progressively increasing regolith thickness:

$$\text{for } C > C_c, \quad W = W_c k e^{-k(C-C_c)}$$

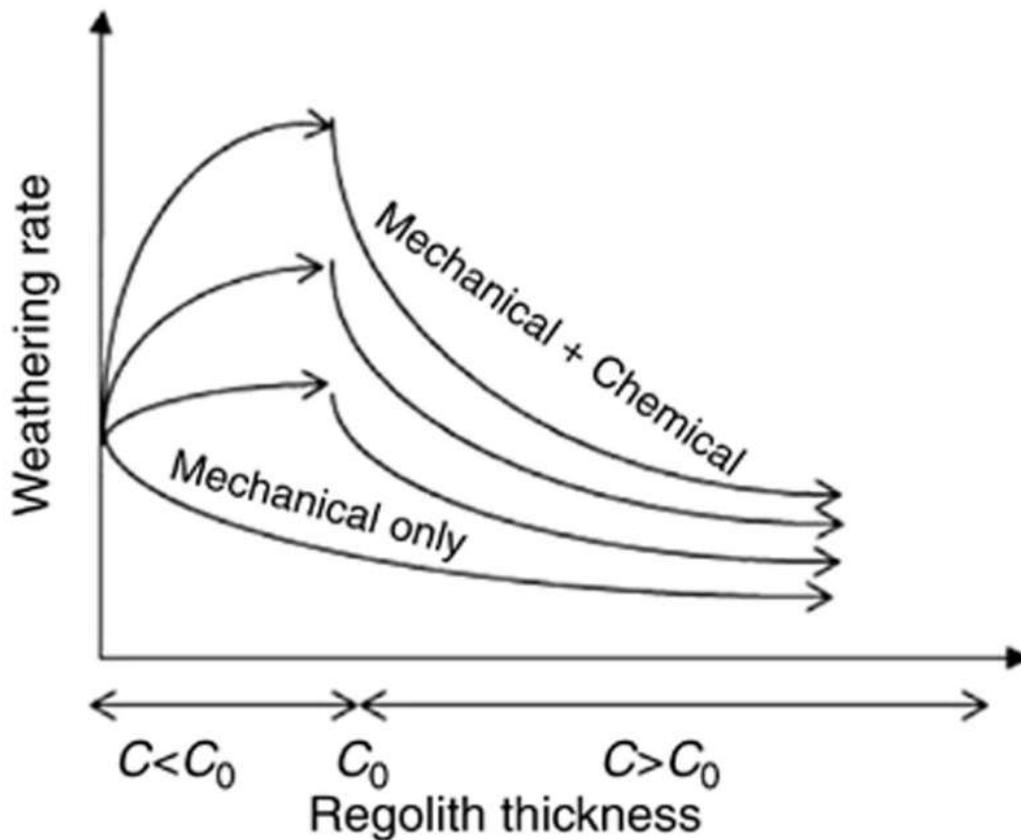


Figure 4.5 Ahnert's (1987) concept of regolith thickness as a function of weathering rate and type of weathering.

This is a negative exponential function, in which the unit-weathering rate decreases with increasing regolith thickness. Beyond the critical thickness, it becomes more likely that weathering agents lose efficacy (for instance, become saturated) and also more likely that the regolith components approach thermodynamic equilibrium.

The regolith thickness model of Ahnert (1987) , as well as Phillips' (2014) concept of convergent weathering intensity and weathering rate, could be applied to explain the "unexpected" regolith thickness differences observed by Bazilevskaya et al. (2013) . In their study in the Virginia (USA) piedmont, granitic profiles were much thicker than those on diabase, which is the opposite of what was expected, based on mineralogical composition. Although potentially weathering is slower due to higher silica content, the granitic rock was deeply permeated through the loss of biotite, and was able to develop a critical depth thicker than that of nominally faster-weathering diabase. Migoń (2013) summarized a number of factors that are conducive to deep weathering mantles, reproduced in Table 4.3 .

**Table 4.3****Migoñ's (2013) Summary of Factors of Enhancement or Limitation to Deep Weathering Mantles**

Rock property factors	Enhanced deep weathering	Limited deep weathering
Mineralogy and rock chemistry	<ul style="list-style-type: none"> <li>• Rocks with a large proportion of easily weatherable minerals</li> <li>• Rock with less Si, but more Fe, Ca, Na, and Mg</li> </ul>	<ul style="list-style-type: none"> <li>• Quartz-rich rocks</li> <li>• Rocks higher in Si, or rich in feldspars</li> </ul>
Fabric	<ul style="list-style-type: none"> <li>• Weak fabric, susceptible to granular disintegration</li> <li>• High primary or secondary porosity</li> </ul>	<ul style="list-style-type: none"> <li>• Interlocking crystal structure</li> <li>• Low primary porosity</li> </ul>
Discontinuities	<ul style="list-style-type: none"> <li>• High fracture density</li> <li>• Discontinuous, irregular joint network which slows water movement</li> </ul>	<ul style="list-style-type: none"> <li>• Massive rocks with widely spaced fractures</li> <li>• Wide and continuous joints, rapid through-flow</li> </ul>

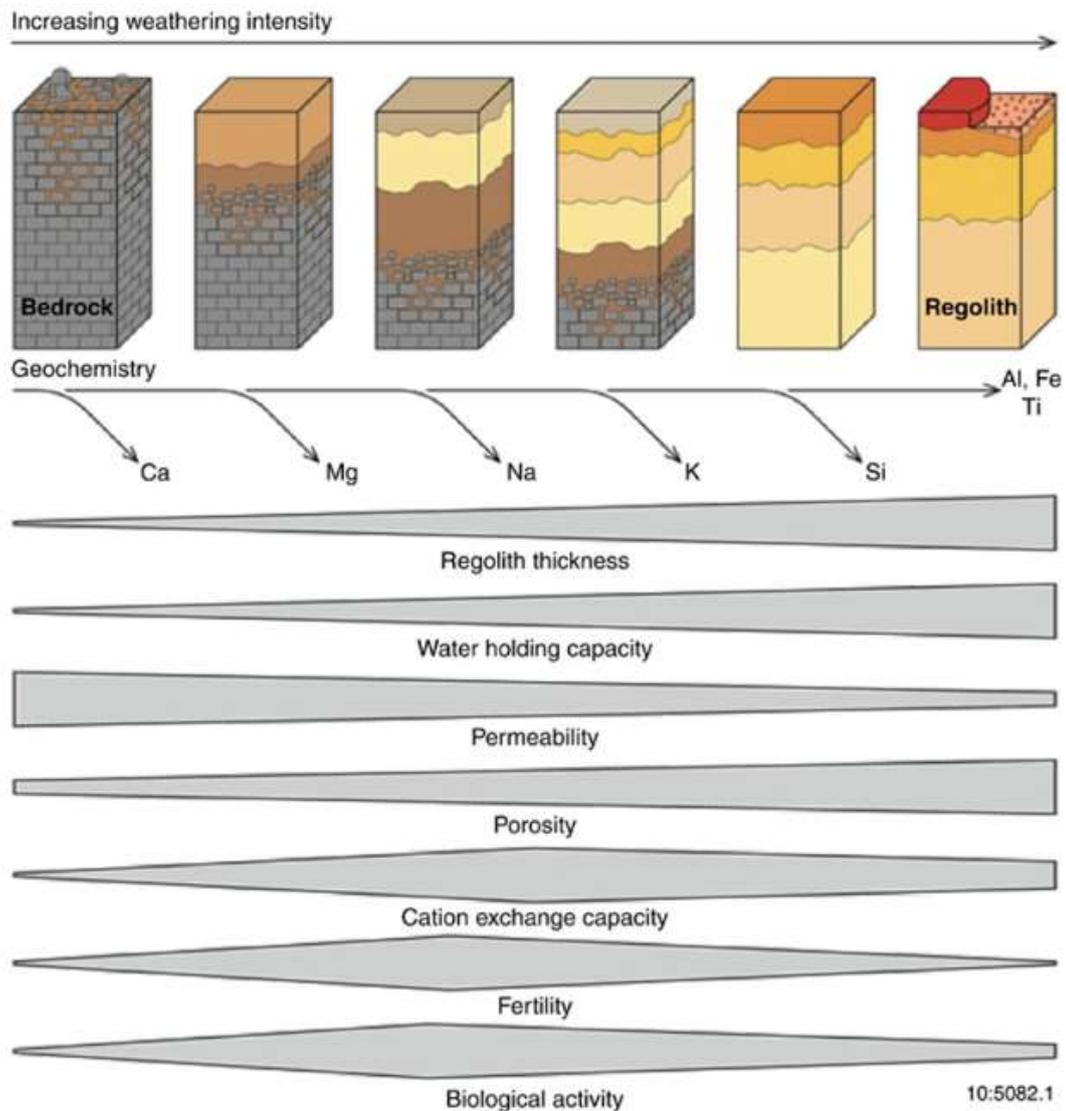
With long term decay of the regolith comes mass loss. Although Ahnert's (1987) model confines itself simply to weathering production, the impacts of surface material loss also determine regolith thickness. The concept of "weathering-limited" versus "transport-limited" landscapes comes into play. Transport-limited landscapes are those in which erosion processes are slower than weathering production. Tropical conditions of high moisture and temperature are commonly used to illustrate transport limitation. Weathering-limited is applied to locations where erosion exceeds weathering production. Textbooks cite deserts and polar areas as lacking sufficient weathering to keep pace with erosion. Pope (2013b) and Pope et al. (1995) pointed out that weathering efficacy is not the sole factor in weathering/erosion limitation. Factors that retard erosion, such as dense vegetation, lower slopes, and higher infiltration rates tend to be associated with many presumed "high weathering" climates. Conversely, slopes lacking vegetation and with higher rates of surface runoff would result in high erosion, regardless of weathering intensity.

Lastly, chemical denudation requires mention in regolith thickness. Weathering removes material by dissolved ions in ground- and soil-water. Volume is not necessarily lost: deeply decayed saprolite can be thought of as a "skeletal" remnant of former rock, with as much as one-third of the mass missing in the form of removed ions and translocated fines ( Schaetzl and Anderson, 2005 , p. 172). In other cases, mass loss results in a compaction of the surface, and sometimes lowering. Pavich (1989) attributed the undulating geomorphic surface of the Appalachian Piedmont to long-term decay and then settling of the resulting saprolite, as opposed to an old erosion surface (peneplain). This process may account for other deeply decayed and geomorphically old surfaces elsewhere ( Phillips, 2002 ).

#### 4.6. Rock decay in the fourth dimension: time and rates in the Critical Zone

Weathering rates have been studied from the scale of mineral lattices to soil profiles to entire landmasses. Although weathering varies by three-dimensional space, across geographic area and into the surface thickness, it also varies over time. Observing weathering over time introduces the fourth dimension, and the ability to classify rates of change done by weathering. Figure 4.6, based on the work of Wilford (2011b), illustrates the idealized evolution of regolith, with the concomitant losses of cations (useful for calculating weathering ratios), as well as changes in properties of

the regolith such as permeability and water holding capacity, fertility, and biological activity.



In chemical decay processes, the relationship between reaction rate and temperature is demonstrated by the Arrhenius equation:

$$\ln k = \ln A - \ln \left( \frac{E_a}{RT} \right)$$

where  $k$  is the rate constant,  $A$  is the molecular collision factor,  $E_a$  is the activation energy of the substance,  $R$  is the gas constant, and  $T$  is the temperature. In low temperature geochemical reactions such as rock decay, the nature of the rock composition factors into the activation energy, the efficacy of weathering agents

would be reflected in the molecular collision factor, and temperature would become an important variable, the relation of rock decay to climate. Thus, rates would increase when ambient temperatures increase and/or decay agents increase. The equation can be solved for activation energy  $E_a$  when differing decay outcomes (such as mass loss,  $\Delta M$ ) can be compared at different temperatures (such as between locations, or the same location at different times):

$$E_a = \frac{R(\ln(\Delta M_1 / \Delta M_2))}{T_1^{-1} - T_2^{-1}}$$

In this case, it can be seen how the activation energy, the minimal energy needed for the reaction to take place, is lower when ambient temperatures are higher.

Contact time with water is important to chemical reactions in the regolith. The Critical Zone encompasses the vadose profile, in which the matrix is water-unsaturated except for specific saturation events or seasons. The amount of water in contact with soil or regolith particles ranges from adsorbed films on particles (“hygroscopic” water) to water filling small pores (“capillary” water) to saturated conditions in which pores are completely full, and varies considerably based on texture/particle size and composition (Schaetzl and Anderson, 2005). Saturated water is influenced by positive gravitational potential, whereas pore and adsorbed water would have negative potential (osmotic or matric, respectively), thus greater attraction to the particle (Brady and Weil, 2008). Adsorbed water is theoretically capable of chemical decay processes on mineral surfaces (Zilberbrand, 1999). However, without resupply these thin film solutions acquire more negative potential, and would quickly saturate and lose effectiveness. Studies of mineral decay associated with adsorbed or hygroscopic films are mostly limited to the stone conservation literature (Camuffo, 1995; Larsen and Nielsen, 1990), with relatively little attention to natural environments, although the status of nanoscale observation is well developed (MacInnis and Brantley, 1993; Brantley et al., 1993; Dorn et al., 2013b).

In general, the residence time ( $t_R$ ) of water in the soil and regolith is a function of both volume ( $V$ ) and flow rate ( $Q$ ) (Langmuir, 1997):

$$t_R = \frac{V}{Q}$$

Residence time for water in the soil increases with increasing depth in the soil and regolith. Within the soil profile, depending on texture and permeability, residence times of water are on the order of <1–4 months, whereas deeper into the regolith, residence times are measured in years to decades (Bleam, 2012; Matsutani et al., 1993; Langmuir, 1997; Stewart and McDonnell, 1991). Some mineral-water reactions take place nearly instantly, but reactions such as hydrolysis and recrystallization may evolve over many thousands of years (Langmuir, 1997). Thus, longer exposure to water results in greater potential for decay, up to the point of solute saturation in the water. As an illustration of contact time correlation to chemical activity, lake

alkalinity has been shown to increase with time as the groundwater residence time increases in the soil and regolith (Wolock et al., 1989). Although chemical reaction rates would theoretically slow as solutions become saturated or decay products reach end-stage equilibrium, it appears that long-resident (25 years) regolith water does not show a rate decrease (Wolock et al., 1989).

At the mineral boundary layer, surface area and solution concentration are relevant to the decay rate. Lasaga (1994) modeled these relationships as follows, with a given mineral  $\theta$  influenced by a solution  $i$  :

$$r = \frac{dc_i}{dt} |_{\text{diss}} = \frac{A_\theta}{V} \beta_{i\theta} k_\theta$$

where  $dc_i / dt$  is the change in solution concentration as the dissolution reaction progresses,  $A_\theta$  is the exposed area of the mineral,  $V$  is the volume of fluid in contact,  $\beta_{i\theta}$  is the stoichiometric constant, and  $k_\theta$  is the rate constant. Bricker et al. (2005) indicate that difficulties in estimating exposed surface area or “wetted surface area” (which also includes internal microfractures) would be cause for discrepancies in modeled versus field-observed dissolution rates. However, Pope (1995) demonstrated a means to quantify internal weathering, by image processing of backscatter electron micrographs that would be applicable to the model.

Individual minerals have different susceptibility to decay, so would have different rates of decay. A given rock (composed of several minerals) would decay initially as fast as its most susceptible minerals, although its remnants would comprise its least susceptible minerals. Lasaga et al. (1994) calculated the persistence of various minerals subjected to decay:

$$\frac{dR}{dt} = -k_i \bar{V}_i$$

where for a specific mineral grain  $i$  the change in the mineral grain radius over time is a product of the dissolution rate  $k$  and the molar volume  $\bar{V}$  of the mineral grain. The results mirror those assumed by Goldich’s (1938) sequence of mineral weathering. A single 1-mm crystal of quartz would persist 34 million years at a log rate of  $-13.39$  mol/m<sup>2</sup> s. Rock decay end-product kaolinite would persist six million years at a log decay rate of  $-13.28$  mol/m<sup>2</sup> s. At the other end of the spectrum, a 1-mm crystal of anorthite (calcium feldspar) would last just 112 years at a log decay rate of  $-8.55$  mol/m<sup>2</sup> s, and the pyroxene wollastonite (calcium silicate) persists 79 years at a log decay rate of  $-8.00$  mol/m<sup>2</sup> s. The minerals with low-residence-time and high-decay-susceptibility also corresponds high “weathering potential” (Reiche, 1950). It should be noted that these trends are not absolute. Exceptions are possible if minerals have different susceptibilities to decay agents, as in the case where olivine persists anomalously in microenvironments devoid of organic decay agents (Wasklewicz, 1994). For the most part, mineral persistence explains the prevalence of quartz in the

terrestrial surface environment, and provides a means of measuring the maturity of sediments (Folk, 1951) and the relative age of soils by mineral depletion (Birkeland, 1999). Rates of weathering of individual minerals are covered well in earlier edited works such as Colman and Dethier (1986) and White and Brantley (1995).

Calculating the persistence of a rock, composed of many minerals, becomes more complicated, and the persistence of soils and regolith profiles much more complicated. Further generalizations are required. Although most researchers recognize the primacy of the nanoscale boundary layer as the first occurrence of decay, getting from there to rock, pedon, and regolith scale is not straightforward, with a “gray box” understanding of up-scale connections and factors. All the while, the holistic nature of combined processes and combined factors needs to be recognized. Decay outcomes of soil and regolith production (matched against negative feedbacks of erosion), and contribution to runoff solute loads, provide a sense of the combined processes. Recent studies calculated generalized regolith production rates from an assemblage of individual mineral or rock weathering rates. Ferrier et al. (2010) attempted this method at Rio Icacos, Puerto Rico, arriving at rates of <200 to >3000 moles/hectare/year. Cation and silica depletion across an assemblage of rock types was used by Hartmann and Moosdorf (2011) to estimate chemical flux across multiple watersheds in Japan, potentially extrapolated to global scales.

Rates of soil and regolith formation and residence time have been a perennial topic of research, but with limited ability to pin down definitively, until recently. Researchers relied on relative and semiquantitative dating techniques (such as soil development indexes, weathering rinds on cobbles, depletion of certain minerals, accumulation of caliche, cf. Chapter 8 in Birkeland, 1999). Paleosols and buried soils could sometimes be dated with radiocarbon, luminescence, or isotope decay in volcanic deposits on bracketing strata. However, the understanding of cosmogenic isotope pathways in the Critical Zone (Graly et al., 2010 for a recent study) led to revolutionary advances in discerning soil and regolith evolution. Pavich et al. pioneered the use of Be-10 to estimate ages of soil and regolith profiles, first in the Virginia piedmont, and later elsewhere in the Appalachian Mountains (Markewich et al., 1994; Pavich, 1989; Pavich et al., 1989; 1985), arriving at a regolith residence time approximating one million years. The methods have since been applied to many locales (cf. Heimsath et al., 1997 ; Riebe et al., 2003 ; Hewawasa et al., 2013 ). The residence time of a particular regolith particle increases as it progresses upward to the surface (see Fig. 4.1). Dosseto et al. (2011) , using uranium-series isotopes, found soil residence times less than 10 ka at the lowest reaches of a regolith profile (□16 m), but exceeding 80 ka at the surface.

In the simplest sense, the advance of the rock decay front into fresh bedrock over time can be conceptualized as a crude linear rate, for instance, × meters per million years, presuming a profile of measurable depth and an exposure surface age. Hewawasa et al. (2013) conclude a weathering front advance rate of 6–14 mm/ka in slowly weathering crystalline rock in Sri Lanka, based on a regolith base □10-m deep and ages constrained by cosmogenic nuclide accumulation in the soil as well as river dissolved loads. Rates in reality are most certainly not linear, instead are perhaps more nonequilibrium to a point, but most likely complex and chaotic ( Phillips, 2001b ), not unlike other geomorphic systems (Phillips, 2005 , 2014).

#### 4.7. Conclusions

Rock decay processes are at the center of the Critical Zone, the weathering engine that modifies the Earth's crust to adjust to surface atmospheric, hydrologic, and biotic conditions. It is for this reason that weathering research has reached tremendous output, not only to explain surface processes, but also to contribute toward better understanding of surface and near-surface hydrology, ecology, climate regulation, and applications in natural resources. Weathering is a synergistic mechano-chemical process; individual processes such as stress release, pressure, and water and acid reactions provide positive feedback to each other, and in general work together.

Weathering processes focus on a nanoscale boundary layer of mineral destruction, although nested environments, at increasing scales, are the sources of weathering agents and process modifying factors. But, although weathering takes place at a point location, the integration of weathering throughout the thickness of the regolith provides the key benefit of the Critical Zone concept: a holistic, three-dimensional porous body, leaky to the interior of the surface crust by groundwater, and to the subaerial exterior via organisms and the atmosphere. The formation of this regolith zone is found to be thousands if not millions of years of development. The rates of formation, and of depletion, provide a critical focus of ongoing studies across geographic areas.

4.7.1. A Note on Critical Zone Interdisciplinarity and Weathering System Science

The Weathering System Science Consortium, one of the first scientific collaborations inspired by the NRC's Critical Zone charge, is an ideal concept of converging ideas from different disciplines – geomorphology, hydrology, pedology, mineralogy, petrology, geochemistry – tackling a common research goal. It has been successful enough that the consortium broadened its name, and the CZEN now spreads into other sciences such as ecology and geohydrology. Yet, the Critical Zone movement has not reached all in the broad scope of weathering system science, nor has CZEN been able to recognize parallel (and in some cases long-term) efforts toward the same goals. Case in point, there are still “schools” of research that continue to talk past each other without much cross fertilization. This lack of cross-fertilization is best recognized in the keystone texts and seminal research papers. Powerful sources that they are, the works of Drever (2005) and Anderson and Anderson (2010), of Taylor and Eggleton (2001) and Scott and Pain (2008), and of Turkington et al. (2005), and of the contemporary researchers contributing or cited therein, seldom cite each other, although often working in the same types of environments. Partially this derives from approaching from different disciplinary cultures, attending only a comfortable availability of conferences and symposia, and some from differing research goals, such as resource development, environmental quality, or landscape evolution. CZEN attempts to bridge those gaps, and has gained the following, although some by happenstance and less by direct recruiting outside the immediate scholarly connections.

## **Chapter 5**

# **Soil Morphology in the Critical Zone: The Role of Climate, Geology, and Vegetation in Soil Formation in the Critical Zone**

John C. Dixon Department of Geosciences, University of Arkansas, Fayetteville, Arkansas, USA

### **Abstract**

The Critical Zone has been defined by the National Research Council (2001) as the heterogeneous, near-surface environment in which complex interactions involving rock, soil, water, air, and living organisms regulate the natural habitat and determine the availability of life-sustaining resources. The Critical Zone extends from the top of the vegetative canopy to the bottom of the groundwater aquifer. Soil, thus, forms the core of the Critical Zone as it incorporates a substantial part of this entire region of the “skin” of Earth. This chapter examines the nature of soils, their formation and their dependence on interactions, and feedbacks between climate, geology, and vegetation.

Keywords Critical Zone soil formation soil properties soil state factors, soil reddening

### **5.1. Introduction**

Soil lies at the core of the Critical Zone (Fig. 5.1) as it represents the “membrane” across which, or through which water and solutes, energy, gases, solids, and organisms are exchanged between the atmosphere, biosphere, hydrosphere, and lithosphere. Water is the force that drives the majority of these exchanges within the soil as well as in the Critical Zone as a whole (Lin, 2010). Soil is distinguished from regolith by two essential characteristics namely the presence of genetically related soil horizons and the presence of living and dead organic matter. Yet simultaneously, regolith is an essential component of the soil profile and the Critical Zone as a whole as it is either the lithified or unconsolidated parent material from which the soil is derived. Soil property variability within the Critical Zone is a function of the interaction of five fundamental factors: climate, organic activity, relief, parent material, and time (Jenny, 1941). This chapter focuses on the fundamental roles played by climate, vegetation, and geology (i.e., relief and parent-material) on the formation and associated properties of soil developed in the Critical Zone.

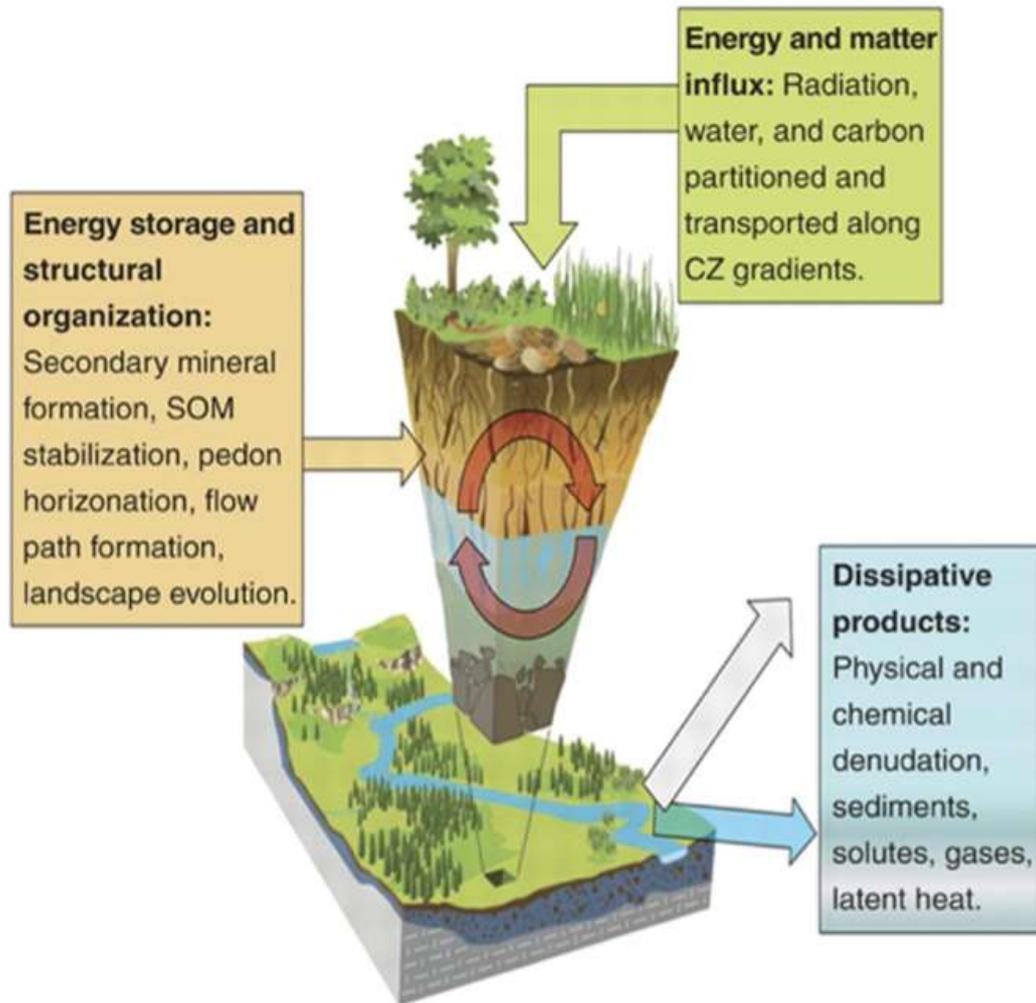


Figure 5.1 Conceptual model of the Critical Zone showing the prominent position of the soil.

### 5.1.1. The Nature of Soil

Soil is a natural body comprised of solids (i.e., minerals and organic matter), liquid, and gases that occur at, or near, the land surface, and is characterized by one or both of the following: horizons, or layers, that are distinguishable from the initial formative material as a result of additions, losses, transfers, and transformations of energy and matter, and the ability to support rooted plants in a natural environment (Soil Survey Staff, 1994).

The upper limit of soil is the boundary between soil and air, shallow water, live plants, or plant materials that have not begun to decompose. Areas are not considered to have soil if the surface is permanently covered with water too deep for the growth of rooted plants. The lower boundary that separates soil from the nonsoil underneath is more difficult to define. Soil consists of horizons near the surface of the Earth that, in contrast to the underlying parent material, have been altered by the interactions of climate, relief, and living organisms over time. Commonly, soil grades at its lower boundary to hard rock or to unconsolidated materials virtually devoid of animals, roots, or other indications of biological activity. For purposes of classification, the lower boundary of soil is arbitrarily set at 200 cm (Soil Survey Staff, 1994).

Soil can be viewed as the ultimate product of weathering of rocks and minerals at or near the surface of the Earth. It represents an equilibrium weathering product in which primary minerals are in metastable equilibrium with temperature, pressure, moisture, and gaseous conditions in Earth environments. As environmental conditions change however, they may be accompanied by responses in the weathering system. Weathering transformation of rock and regolith are examined in Chapter 4 by Pope.

### **5.1.2. The Soil Profile**

Soils consist of a number of readily identified master horizons that reflect varying degrees of alteration of an original parent material (Fig. 5.2). Soils develop as a consequence of the biophysical and biogeochemical weathering of a parent material. If the parent material is a lithified or crystallized rock material, the associated soil horizon is referred to as the “R” horizon, and defined as the consolidated bedrock underlying the soil ( Birkeland, 1999 ). This horizon is by definition, one of fresh unaltered rock material in which weathering processes have yet to affect alteration, except for perhaps some oxidation. Above the “R” horizon is the “C” horizon. This horizon is frequently referred to as the regolith horizon and consists of material little affected by pedogenic processes, but displaying strong weathering. It lacks the properties of the overlying, truly pedogenic horizons. Most of the “C” horizons consist predominantly of mineral matter that may be like or unlike the material from which the overlying pedogenic horizons have been presumed to have formed (Soil Survey Staff, 1994). Where soil parent materials are unconsolidated they are most commonly designated as the “C” horizon. However their designation as a “D” horizon has been proposed by Tandarich et al. (1994) , but this designation was never formally adopted. Despite this, some Critical Zone researchers do recognize a “D” horizon, as well as horizons transitional between “C” and “D” ( Lin, 2010 ; Brantley et al., 2006 ; 2007 ).

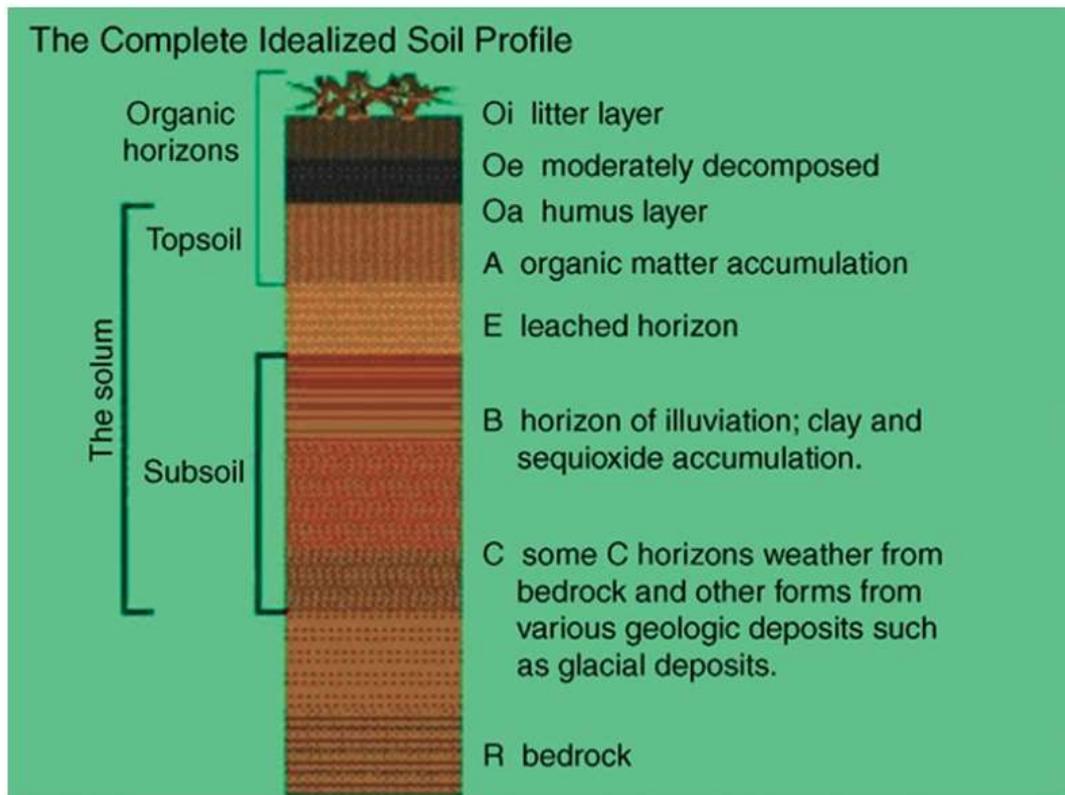


Figure 5.2 Idealized soil profile showing the dominant master horizons present in well-developed soils of the middle latitudes.

Subsurface horizons that are so strongly impregnated with calcium carbonate that their morphology is determined by the carbonate are designated as “K” horizons (Gile et al., 1965). Calcium carbonate coats or engulfs nearly all primary grains in a continuous layer. The uppermost part of a strongly developed “K” horizon is laminated, brecciated, and/or pisolitic (Machette, 1985; Dixon and McLaren, 2009). Continuous layers of strongly cemented carbonate are referred to as hardpan calcrete (Dixon and McLaren, 2009).

Above the “C” or “K” horizon the first of the truly pedogenic soil horizons occur. The “B” horizon is characterized by the obliteration of all, or most of the original parent material structure, and shows one or more of the following characteristics:

1. The concentration of illuvial clay, iron, aluminum, organic matter carbonates, gypsum, or silica in combination or alone.
2. Concentration of residual sesquioxides.
3. Sesquioxide coatings that make the horizon conspicuously lower in color value, higher in chroma, or redder in hue without apparent iron illuviation.
4. Alteration that forms secondary silicate clays or oxides, or both, and that forms a granular, blocky, or prismatic structure if volume changes accompany changes in moisture content.
5. Brittleness or strong gleying ( Soil Survey Staff, 1994 ).

Above the “B” horizon, and below the “A” horizon is the “E” horizon. This horizon is a mineral horizon distinguished by the loss of silicate clay, iron, aluminum, or some combination of these constituents leaving a concentration of sand and silt

particles. This horizon exhibits obliteration of all or most of the original parent material structure.

The uppermost mineral horizon of a soil is designated as the “A” horizon. This is a horizon that has formed at the surface of Earth or below an organic matter-dominated horizon. It exhibits obliteration of all or much of the original parent-material structure and exhibits one or both of the following:

1. An accumulation of humified organic matter strongly mixed with the mineral fraction and not dominated by properties characteristic of “E” or “B” horizons.

2. Properties resulting from cultivation, pasturing, or similar kinds of disturbances (Soil Survey Staff, 1994).

Surface horizons dominated by organic matter are designated as “O” horizons. These horizons may possess a mineral matter component, but it is generally substantially less than 50% by weight. “O” horizons are susceptible to water logging, but are now typically drained. Some have never been saturated. The organic component consists of both living and dead floral and faunal components. Dead material is commonly in varying stages of decomposition ( Soil Survey Staff, 1994 ).

## **5.2. Models of soil formation**

Soil formation fundamentally involves the transformation of parent material to materials that are in some kind of equilibrium with the surface of Earth and near surface environmental conditions. Understanding how these transformations occur has been approached in a variety of different ways. This section presents some of the more common approaches that have been used to more fully understand how soils form and change over time, and how soil-forming factors interact to produce the variability in soil properties observed in the Critical Zone.

### **5.2.1. State Factor Model**

One of the longest standing models of soil formation is that formulated by Jenny (1941) . This fundamentally factorial model, fashioned after the pioneering work of Dokuchev (1883) , identifies five state factors that interact to explain the natural variability in soil properties at different spatial scales. The five state factors are climate, organic activity, relief, parent material, and time.

Climate fundamentally drives rates of parent-material breakdown and biological processes through its control on temperature and moisture at various temporal and spatial scales. The biotic factor as originally defined by Jenny referred to the potential natural vegetation at a site. It referred only to biochemical influences of plants and specifically omitted the role of fauna in soil formation (Jenny, 1958). The relief, or topography factor relates to the configuration of the landscape and fundamentally influences the pattern of water movement across the landscape. In detail, it relates to elements such as slope angle, concavity and convexity, and aspect (Jenny, 1980). The parent material factor encompasses the physical and chemical characteristics of the material from which a soil is derived. It has profound influences on the fundamental physical and chemical properties of the resulting soil. The time factor refers to the length of time that a soil has been developing since the parent material was exposed to surface and near-surface environmental conditions. Fundamentally, it influences the degree of soil development, as expressed in the complexity of soil horizon development.

### **5.2.2. Process-Systems Model**

By the end of the 1950s, models of soil development were beginning to shift from a functional-factorial approach to more of a process-oriented approach (Simonson, 1978; Schaetzl and Anderson, 2005). In the process-systems model, soil formation is envisioned to consist of two essential steps: (1) accumulation of parent material; and (2) differentiation of parent material into soil horizons (Yaalon, 1975). The focus of the process-systems model is on the second part of these steps. It identifies four process bundles: (1) additions, (2) removals, (3) transformations, and (4) translocations ( Simonson, 1978 ). Additions refer to material added to the soil profile from outside sources and include materials delivered from the atmosphere and the biosphere. Similarly, removals refer to losses from the profile as a whole and refer essentially to dissolved materials transported to the groundwater and materials removed by erosion. Translocations refer to losses and additions between horizons, principally from higher to lower horizons, but also under specific conditions from lower to higher horizons. Transformations refer to alterations of inorganic and organic materials in the soil by weathering and decomposition processes. The four process bundles operate simultaneously in the soil profile. Their character and balance governs the ultimate nature of a soil.

### **5.2.3. Runge's Energy Model**

Runge (1973) developed an energy-based model that is basically a hybrid of the state-factor model and the process-systems model. In this model two of Jenny's state factors: climate and relief become the dominant factors. Operating together they become an intensity factor, which Runge defined as the amount of water available for leaching ( $w$ ). When combined, these two factors determine the potential for water to enter and percolate through the soil profile. Thus, the " $w$ " factor acts as an organizing factor that utilizes gravitational energy to organize the soil profile. These ideas have more recently been refashioned into the concept of pedohydrology in which water becomes the chief influence on soil genesis, evolution, variability and function, and these pedogenic characteristics become the principal influence on patterns of water movement in space and time ( Lin, 2010 ). Additionally, Runge's model combines the parent material and organisms factors into a single organic-matter factor ( $o$ ). Runge's two factors  $w$  and  $o$  operating over time constitute the energy model of soil formation (Schaetzl and Anderson, 2005).

### **5.2.4. Dynamic-Rate Model**

The dynamic-rate model stresses interactions between active and passive vectors of pedogenesis and their change over time. The model is an evolutionary rather than developmental model and is based on the idea that soils evolve along interacting progressive and regressive pathways reflecting interactions between exogenous and endogenous vectors (Johnson et al., 1990). Soil evolution is a function of the interaction of dynamic vectors, which are fundamentally exogenous vectors and passive vectors, which are principally endogenous vectors. The dynamic vectors function very similarly to Runge's intensity factor.

The principal dynamic vector categories include energy fluxes, (insolation, heat transfers, entropy transfers, oxidation, and gravity), Mass fluxes (water, gases, and solids), frequency of wetting and drying, organic activity and pedoturbation. The principal passive vectors include parent material, soil chemical environment, water table environment, landscape stability, and pedogenic additions (Johnson et al., 1990).

Soil evolution is the product of interactions between the dynamic (D) and passive (P) vectors and the change in these vectors over time. Intimately associated with the dynamic-rate model are two pedogenic pathways: progressive and regressive pedogenesis (Johnson and Watson-Stegner, 1987). The dynamic-rate model assumes progressive pedogenesis during which there are progressive changes in parent material from a state of simplicity and disorder to a state of increasing order, physiochemical stability, and profile complexity. If pedogenesis subsequently shifts toward simplification, physicochemical instability, and/or shallower profiles, regressive pedogenesis dominates.

#### **5.2.5. Effective Energy and Mass Transfer Models**

Most recently, workers seeking to understand the structure and functioning of the Critical Zone have turned to integrating open-system and thermodynamic theory. This integrated approach investigates the Critical Zone (e.g., especially the soil component) as a system open to energy and mass fluxes forced by radiant, geochemical, and elevational gradients (Rasmussen et al., 2011). This approach uses rates of effective energy and mass transfer to quantify various flux-radiant relations. In particular, it permits the assessment of the relative importance of solar radiation, water, carbon, and physical/chemical denudation mass fluxes to the Critical Zone (soil) energy balance (Rasmussen et al., 2011).

This approach is in many ways an extension of the Runge (1973) model. The model predicts that structural organization within the Critical Zone, particularly soil horizonation, secondary clay mineral and colloid formation, preferential flow path development, vegetation distribution, and landscape evolution, evolves at the expense of increased entropy associated with dissipation of weathering products and gas and latent-heat fluxes (Charover et al., 2011).

### **5.3. Geochemistry and soil development**

Soil formation fundamentally involves the modification of some parent material, either solid bedrock, or transported regolith, by the reaction of water and contained atmospherically and lithospherically derived gases, and the minerals present in the parent material. The “geochemical engine” operating in the soil to move materials around and evolve the soil profile consists of four fundamental components: additions, translocations, transformations, and removals (Simonson, 1978; Brantley et al., 2007; Dixon, 2013). Addition processes involve the contribution of salts, organic compounds, and mineralogenic dust from the atmosphere to the regolith and evolving soil where they are either rapidly transported to depth or accumulate.

Chemically, translocations involve the movement of chemical constituents down through the developing soil profile. Soluble salts are readily dissolved and transported vertically under the influence of gravity while more resistant chemical elements bound in aluminosilicate and other mineral groups may remain near the surface for extended periods of time until they can be released from the host minerals and transported by the process of chelation. Although translocations are fundamentally downward, under favorable circumstances they may occur vertically upward. In either case once the depth of water percolation or capillarity is reached, chemical constituents will be deposited. This results in the formation of a variety of salt-dominated soil horizons (e.g., calcic, sodic, gypsic, sesquioxide horizons) or surface crusts.

Transformations involve the alteration of primary rock-forming minerals and organic matter by a variety of chemical weathering processes (see Chapter 4 for a

detailed discussion) including dissolution, oxidation, carbonation, hydration, and a variety of ligand reactions (Dixon, 2004). The result of these transformation reactions is the formation of new, secondary (principally clay) minerals as well as the release of soluble chemical components, which are either uptaken by vegetation, redeposited as chemical horizons, or removed completely from the soil profile. Transformations are responsible for the formation of clay-rich and salt depleted horizons. Their specific location within a soil profile is fundamentally determined by water table position (Anderson et al., 2007).

Removals collectively refer to those processes that completely remove geochemical constituents from the soil profile. They principally involve the leaching, or complete removal, of soluble materials from the soil profile as well as the removal of chemical constituents as a result of soil surface erosion (Brantley et al., 2007; Dixon, 2013).

### **5.3.1. Soils and Climate**

In the majority of soil formation models, climate assumes premier position and continues to be viewed as a dominant factor in soil formation (Birkeland, 1999). The two most important components of climate are temperature and precipitation, as together they fundamentally control rates of chemical (and to some degree physical) weathering, rates of biomass production and decomposition, and the rates of translocation of solids and dissolved materials. Fundamentally, climate represents the overwhelmingly dominant direct and indirect source of energy in the soil system. The principal soil properties that correlate with climate include organic-matter content, clay content, clay mineralogy, iron mineralogy and color, and patterns of accumulation of calcium carbonate and soluble salts (Birkeland, 1984, 1999).

#### **5.3.1.1. Climate and Organic Matter Content**

Many different measures of organic matter contents in soils are to be found in the literature, including carbon content (Schlesinger, 1990), nitrogen content (Jenny, 1980), and C/N ratios (Post et al., 1985). A global ranking of soil organic matter content (kg/m<sup>2</sup> to 1-m depth) by soil order reveals the following sequence:

Histosols >> Andisols > Inceptisols = Spodosols = Mollisols > Oxisols = Entisols = Ultisols > Alfisols = Vertisols > Aridisols (Eswaran et al., 1993). From quite contrasting geographic settings, Jenny (1961) identified general trends in soil carbon and nitrogen. He found that soil nitrogen increased logarithmically with increasing moisture and decreased exponentially with increasing temperature. For soils from India, he found that organic carbon increased with increasing precipitation. Global patterns of soil nitrogen abundance and C/N ratios closely follow world life zones. Grassland soil C/N ratios range from 10 to 15, whereas desert soil C/N ratios are similar or occasionally lower. This pattern is a reflection of low precipitation and infiltration in desert environments. Many forest and tundra soils display values of 10–20 and are frequently >20 as a result of the toughness of some leaves and needles and low mean annual temperatures (MATs). Low-latitude forests commonly display C/N ratios similar to desert and grassland soils as a result of abundant leaf fall and rapid rates of decomposition (Birkeland, 1999).

#### **5.3.1.2. Climate and Clay Content**

From first principles, it would be reasonable to expect that a linear relationship exists between clay content of soils and climate. One might expect that as temperature and precipitation increase, the clay content of soils would also increase. However,

many factors complicate such a simple relationship. The relatively recent identification of aeolian dust as a substantial contributor to clay accumulation in all climates, parent-material lithology, and topography all play important roles. However, despite these complex influences, several studies exist that establish the fundamental nature of the climatic relationship. Jenny (1935) was one of the earliest to establish the nature of the climate–clay relationship. He showed that a linear relationship existed between moisture and an exponential relationship with temperature. Sherman (1952) working in Hawaii demonstrated an overall increase in clay abundance with increasing precipitation. Strakhov (1967) demonstrated progressive accumulation of weathering-derived clay with progressive increases in temperature and moisture. Webb et al. (1986), in a climatic gradient from 64 to 200 mean annual precipitation (MAP) reported subsoil clay contents doubling. In a comprehensive N–S and E–W pair of climatic transects, Ruhe (1984a and b) reported contradictory trends in soil properties with increasing moisture and increasing temperature, respectively. These contradictory trends he ascribed principally to the effect of parent-material influences on clay formation.

### **5.3.1.3. Climate and Clay Mineralogy**

Weathering in the Critical Zone is discussed in detail in Chapter 4 by Pope, so will only be explored here briefly. Climate strongly influences soil–clay mineralogy (Curtis, 1990). Primary rock-forming minerals are fundamentally unstable under Earth surface and near-surface conditions. Here moisture, temperature, gaseous, and pressure environments are markedly different from those that prevail where mineral formation takes place. Consequently, minerals transform into secondary products (i.e., generally clays) that are more kinetically stable under prevailing conditions. These transformations are driven by reactions of water, oxygen, and carbon dioxide with mineral surfaces. Rates of transformation are fundamentally controlled by temperature and moisture. In a classic study by White and Blum (1995) of 68 watersheds worldwide they concluded that chemical weathering rates are controlled by precipitation for any given temperature, but the strength of the precipitation correlation increases exponentially with increasing temperature. Accelerated chemical weathering occurs in warm, wet environments compared to cool, wet environments. This, however, does not mean that chemical weathering and, thus, soil formation is inhibited in cold climates (Dixon and Thorn, 2005). The nature of the secondary weathering products varies markedly with climate, principally in response to moisture availability. Generally speaking, warm, humid climates are dominated by kaolinite; cool, humid environments are dominated by smectite and montmorillonite; whereas, sepiolite and palygorskite dominate in arid environments (Birkeland, 1999).

Tardy et al. (1973) demonstrated a change from smectite/gypsum-dominated crest slopes and smectite/gypsum/halite-dominated toe slopes on soil catenas in arid environments; through kaolinite and Fe-hydroxides-dominated crest slopes and smectites and carbonates in toe slope soils in semiarid environments; to gibbsite and Fe-hydroxide-dominated crest slopes and gibbsite-dominated toe slopes in humid environments (Fig. 5.3). Even under constantly humid (wet) environmental conditions, specific clay minerals dominate, depending on the precise amounts of precipitation received (Sherman, 1952).

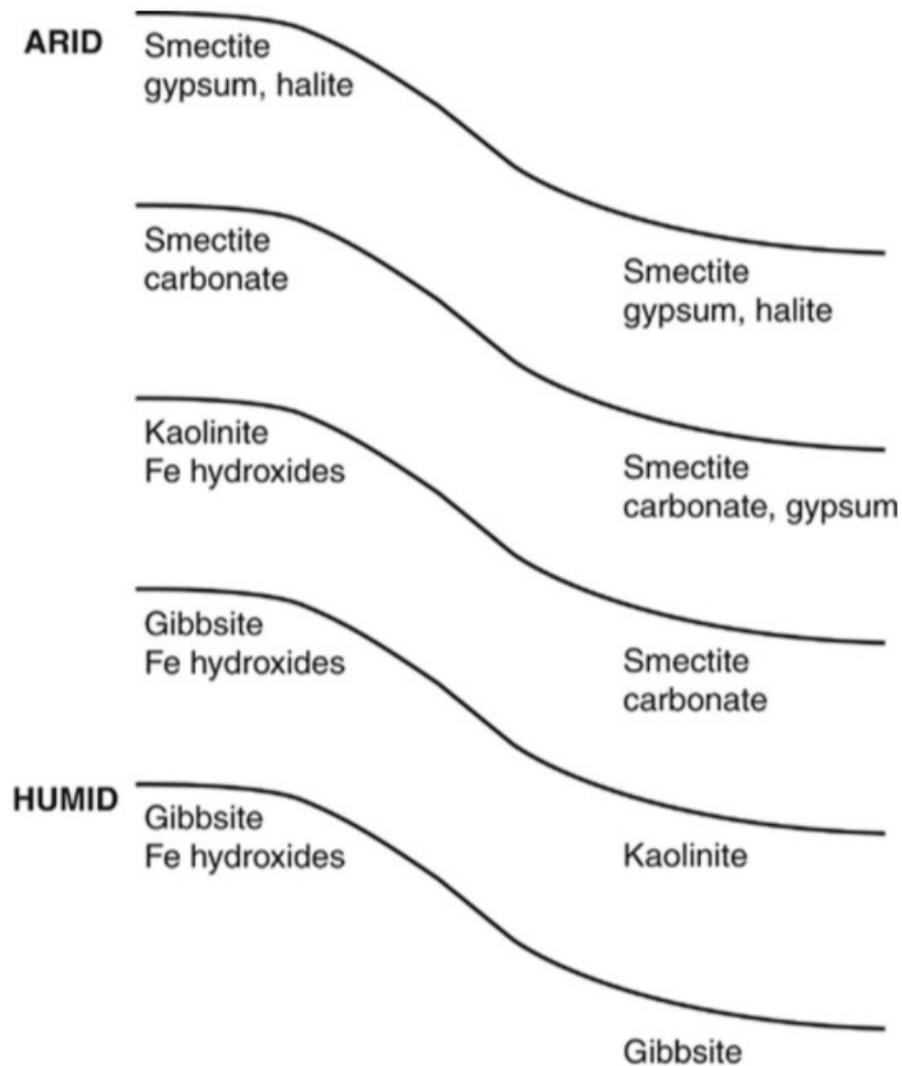


Figure 5.3 Schematic of the effects of climate on clay mineral formation and distribution on a generalized slope.

McFadden (1988) working in the southwestern United States, investigated the nature of clay–mineral assemblages along an arid-semiarid-xeric transect. At the arid end of the transect, the dominant minerals are (i.e., in relative abundance) illite-smectite-palygorskite-kaolinite, whereas at the xeric end of the transect, the sequence is kaolinite-vermiculite-illite-smectite.

More recently, Egli et al. (2003) conducted a study of a soil climosequence in the Italian Alps, along a strong precipitation and temperature gradient. They found that weathering rates and the formation of soil smectites decreased both with increasing and decreasing elevation, and that this strongly nonlinear pattern was intimately related to pronounced podsolization at treeline where higher precipitation rates and the production of chelates promote smectite formation. This study amplifies the strong role vegetation plays in soil–clay mineral formation in conjunction with climate. Vegetation serves to bind soil particles and to reduce erosion, thus, increasing reaction time between soil solutes and mineral particles and completeness of mineral dissolution. In addition it provided  $H^+$  for acid reactions while photosynthesis and subsequent degradation pump  $CO_2$  through the soil, which facilitates additional chemical reactions (Curtis, 1990).

#### **5.3.1.4. Soil Reddening**

Soil color, especially soil redness is strongly correlated with climate (Birkeland, 1999). It appears that temperature exerts a stronger control than precipitation on redness as illustrated by the strong red colors of the southeastern United States, the humid tropics and hot deserts of the southwestern United States and Baja California (Birkeland, 1999). In contrast, soils are less red in cooler environments, such as those along the United States/Canadian border and Arctic and alpine environments (Birkeland, 1999). In a study of soil redness and iron mineralogy in a transect across the European Alps, Schwertmann et al. (1982) found that soil color reddens with increasing MAT and decreasing MAP. Iron mineralogy, as expressed by the ratio hematite:hematite + goethite, shows a strongly similar trend. A similar study by the authors in Brazil shows a similar trend with respect to mean annual air temperature (Kampf and Schwertmann, 1983).

#### **5.3.1.5. Accumulation of Calcium Carbonate and Soluble Salts**

With increasing aridity, the accumulation of calcium carbonate and other soluble salts begins to occur in soils. In general, carbonate begins to accumulate in soils where MAP is less than 50 cm, whereas gypsum accumulates in soils where MAP is less than 10 cm (Dan and Yaalon, 1983). MAT also influences the accumulation of the more soluble salts. In the mid-continent United States, the boundary between calcic and noncalcic soils is a MAP of 50 cm and a MAT of 5–6°C in the north. In southern Texas the boundary is at approximately 60 cm MAP and 22°C MAT. In the western United States, the climatic conditions associated with carbonate accumulation are radically different where deposition occurs under significantly lower precipitation regimes and leaching is substantially less effective. Here, carbonate accumulates under a MAP of 20 cm and a MAT of 9.5°C (Birkeland, 1999). Precipitation chemistry is also a significant component on patterns of carbonate accumulation.

Depth to the top of the carbonate layer varies directly with precipitation and is a reflection of the average depth of infiltration of moisture. In more arid environments, depths are generally shallower, whereas in wetter environments depths are greater. This pattern is complicated, however, by the fact that parent-material properties strongly influence the patterns of carbonate and salt accumulations (see Dixon and McLaren, 2009 and Dixon, 2009 for detailed discussions).

#### **5.4. Soil properties and geology**

For the purpose of the present discussion, geology is treated as a collective term for both the influence of topography and the influence of parent-material properties on soil development. Topography largely controls the pattern of water movement on slopes through slope form and angle (Ruhe, 1960; 1975; Conacher and Dalrymple, 1977; Schaetzl, 2013), the position of the water table, as well as aspect which integrates topography and climate (Cooper, 1960; Carter and Ciolkosz, 1991; Rech et al., 2001; Egli et al., 2007, 2010). Parent-material exerts strong controls on soil texture and structure, chemistry, color, all related essentially to patterns of water movement and associated transportation of solids and solutes as well as the influence it has on vegetation.

### 5.4.1. Soil Catenas

Soil catenas are associations of soils on a single slope. They display different characteristics depending on slope position, but they are all linked by the common association they share with respect to patterns of water movement on the slope and associated movement of solids, solutes, and gases (Schaetzl, 2013). The principal slope elements of a catena consist of summit, shoulder, backslope, footslope, and toeslope elements (Ruhe, 1960 ; Schaetzl, 2013) ( Fig. 5.4 ).

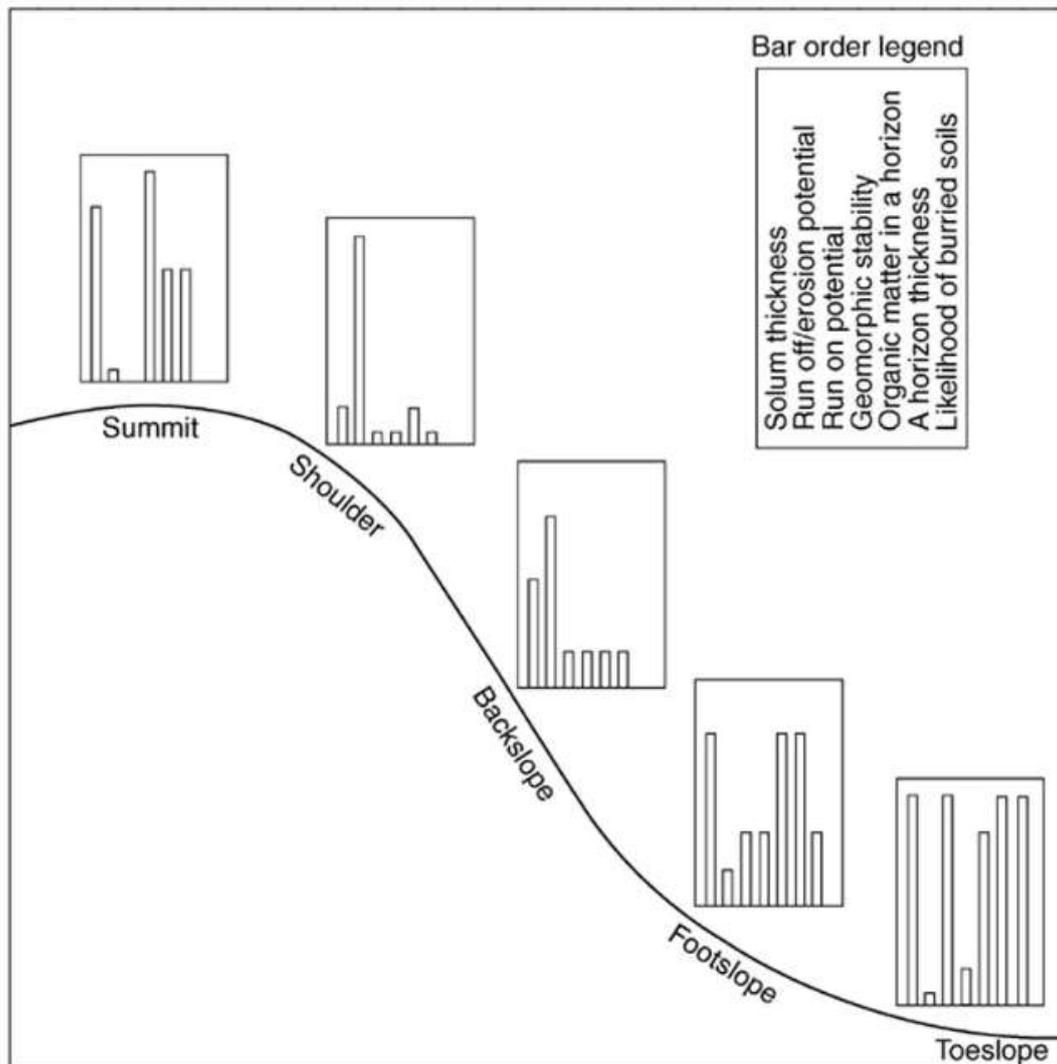


Figure 5.4 Effects of hillslope position on soil properties in a humid climate.

Summit slopes are stable, with minimal erosion. Chemical weathering dominates over physical weathering. Soils are typically thick, as infiltration dominates, resulting in strong leaching and consequently more strongly developed soils than those on slope elements immediately downslope. In addition, they commonly have thick “A” horizons and accumulate organic matter (Ruhe, 1960 ; Schaetzl, 2013).

Shoulder slopes exhibit maximum convexity and, thereby, are dominated by runoff and erosion. Erosion may be so great as to completely outstrip soil formation (Walker and Ruhe, 1968). Soils occurring in this slope position are commonly thin, young, and least stable. They possess thin “A” horizons, lack organic matter

accumulation and are commonly dry. They are unlikely to accumulate soil materials from higher on the slope (Ruhe, 1960 ; Schaetzl, 2013).

Backslopes are relatively steep, straight slopes lacking pronounced convexity or concavity. They are located between upslope erosional slopes and downslope depositional slopes and as such function principally as transportation slopes but are subject to potential erosion. The soils in these positions display considerable variability depending on the length of the slope segment but generally are thicker than those higher in the catena but slightly thinner than those in lower topographic positions. They typically display relatively thin “A” horizons and accumulate relatively small quantities of organic matter (Ruhe, 1960 ; Schaetzl, 2013).

Footslopes are the concave portions of a catena toward the bottom of the slope. This slope element is dominated by accumulation of materials carried from higher on the catena in suspension, solution and wash, both as surface wash and subsurface throughflow. Soils developed in this location are commonly thick because of the accumulation of materials from higher slope positions, as well as relative geomorphic stability, although potential for some erosion still exists. They display the development of thick “A” horizons and accumulation of organic matter. Soils developed in this location also possess the potential for the occurrence of buried soils (Ruhe, 1960 ; Schaetzl, 2013).

Toeslopes occur at the distal ends of footslopes and represent stable, constructional locations at the bottom of catenas. Soils developed in this location are commonly over-thickened as a result of both deposition of materials from higher on the slope, as well as deposition on the slope by fluvial processes. Commonly soils in this position possess very thick “A” horizons and contain multiple buried soils. Because of their inherently wet nature, toeslope soils typically exhibit high plant productivity and organic matter accumulation than at other locations down the catena (Ruhe, 1960 ; Schaetzl, 2013). These locations also can be sites of enhanced chemical weathering. Dixon (1986) noted enhanced clay formation from biotite weathering in these locations in catenas in the Front Range of the Rocky Mountains. The dominant secondary clays forming from biotite weathering were vermiculite and smectite.

The general structure and accompanying soil properties of catenas described above can be broadly applied across landscapes because of the fact that it expresses general patterns of water flow across and through hillslopes. However, in detail, the precise pattern of soil development and accompanying soil properties varies considerably depending on the prevailing climate. Dixon (2013) examined, in detail, the variability observed in catenas developed under a variety of humid, Arctic/alpine, and arid/semiarid climates.

#### **5.4.2. Nine-Unit Slope Model**

Whereas the catenary slope sequence outline can be generally applied at the individual hillslope scale, the nine-unit slope model was developed by Dalrymple et al. (1968) and Conacher and Dalrymple (1977) for the expressed purpose of recognizing that the catena concept should include soil differences that result from variations in drainage, as well as from differential transport of eroded solid and dissolved materials (Milne, 1935 , 1936) ( Fig. 5.5 ). In addition, the nine-unit slope model was developed to incorporate a consideration of soils developed in redeposited materials, hence the evolution of a pedo-geomorphic model. The model is particularly well suited to application in old, stable, deeply incised landscapes. It consists of nine geomorphic units, each with a distinctive set of associated soil properties. The nine units are:

Unit 1. Interfluvial: Chemical weathering and in situ soil development dominate.

Unit 2. Seepage slope: Pedogenically, this unit is characterized by gleying above iron pans or other indurations. Reduced porosity and increased compaction is found in "E" compared to "B" horizons and in upper "B" compared to lower "B" horizon. Mottling, Mn concentrations, and concretions present.

Unit 3. Convex creep slope: Transportation of soil material from upslope by creep and surface and subsurface water movement. Better drained soils than in higher slope positions.

Unit 4. Fall face: Soils shallow to bedrock.

Unit 5. Transportational midslope: Where wash and creep dominate, "A" horizon thickness does not vary by >10% and does not thicken downslope. Where other mass wasting processes dominate, contrasting areas of thick and shallow soils occur.

Unit 6. Colluvial footslope: Heterogeneous soil mantle containing additions from upslope.

Unit 7. Alluvial toeslope: Addition of alluvium. Generally poorly drained soils with thick "A" horizons.

Unit 8. Stream channel wall: Pedogenesis is limited to entisol formation.

Unit 9. Stream channel: No pedogenesis.

### 5.4.3. Soils and Aspect

Soils developed on essentially north-facing slopes in the northern hemisphere or south-facing slopes in the southern hemisphere typically exhibit lower temperature regimes and greater moisture regimes than soils developed on south- and north-facing slopes, respectively. This contrast in thermal and moisture regimes has substantial impact on the properties exhibited by soils and also on the nature of the vegetation associated with such slope contrasts. The principal control being a climatic one in which effective moisture becomes the dominant factor (Cooper, 1960; Finney et al., 1962; Klemmedson, 1964; Macyk et al., 1978; Hunckler and Schaetzl, 1997; Rech et al., 2001).

Generally, soils developed on southerly-facing slopes experience more xeric conditions than those on north-facing slopes and consequently are less strongly developed. Soil horizons tend to be thinner, soil organic carbon tends to be lower, and because of less moisture, parent materials are less weathered (Fig. 5.6). Despite these generalities great variations also occur on specific hillslopes depending on the influence of the other soil-forming factors. Figure 5.6 Two examples of the effect of aspect on soil profile complexity from Michigan and Ohio. Note the greater profile complexity, particularly as it relates to weathering and horizon thickness on north-facing slopes. (After Schaetzl and Anderson (2005).)

In a study of the effect of slope gradient and aspect on soils developed from sandstone in the Ridge and Valley physiographic province of the Appalachian Mountains in Pennsylvania, Carter and Ciolkosz (1991) examined the effect of aspect on numerous soil properties. This study focused on soils developed on summit and shoulder slopes oriented in SW and NW directions. The authors found that with respect to aspect, solum depth, "B" horizon thickness, clay, Fe and Al indices showed a smaller decrease with increasing slope on the NW slopes than on the SW oriented slopes. In addition, "E" horizons were slightly thicker on the NW-facing slopes than the SW-facing slopes. The authors attributed these aspect differences principally to lower evaporation on the northerly oriented slopes and associated more effective

eluviation and illuviation as a result of greater moisture availability. Recent studies of soil formation in the Italian Alps have examined the influence of aspect on the weathering of clay minerals in soils and the patterns of soil organic-matter accumulation. In their study of clay mineral weathering, Egli et al. (2007) identified greater weathering intensities of clays as reflected in greater smectite abundances in soils developed on north-facing slopes than south-facing slope profiles. In these latter soils, smectite only occurred in surface horizons at the highest elevations. In addition, they noted that clay minerals in north-facing slope profiles underwent ionic substitutions in their octahedral sheets that led to reduction in their layer charge. A follow-up study (Egli et al., 2010) focused principally on the effects of exposure on organic matter and the effects of its degradation on soil chemistry. In this study, the authors found that there was a nonlinear relationship of soil organic carbon content to climate. The greatest amount of organic carbon was found to occur in soils developed at treeline. However, a strong aspect influence occurred as soils on north-facing slopes were found to contain greater amounts of organic matter, as well as a significantly lower degree of humification than south-facing slopes. Soils on north-facing slopes exhibited greater amounts of undecomposed and weakly degraded organic matter than soils on south-facing slopes. This difference was credited to lower temperatures and greater acidity and, thus, slower rates of decomposition. Overall, soils on north-facing slopes were more strongly leached than those on south-facing slopes as a result of easier transportation of fulvic acids and accompanying chelation of Fe and Al.

#### **5.4.4. Soils and Parent Materials**

Parent-material characteristics strongly influence the physical and chemical properties of soils. The principal parent-material properties that influence soil properties are those that control rock weathering: texture, porosity, drainage, and mineral composition (Ollier and Pain, 1996). Their influences are most strongly seen in such fundamental soil properties as color, structure, texture, clay mineralogy, and chemistry. Generally, the influence of parent material on soil properties is most strongly expressed in drier climates and in the early stages of soil development. In more humid environments and later stages of soil development, the influence of the other soil-forming factors tends to overshadow parent-material influences (Ollier and Pain, 1996; Birkeland, 1999).

Typically, light-colored parent materials give rise to light-colored soils, fine-textured parent materials give rise to fine-textured and strongly structured soils and coarse-textured parent materials give rise to coarse-textured and weakly-structured soils. Soil chemistry typically reflects parent-material chemistry, and the mineralogical composition of the parent material strongly controls the nature of the secondary clay species formed in soils (Ollier and Pain, 1996). A strong confounding factor to these general trends is the influence exerted by the addition of aeolian dust to soils, which has profound influences on all of these soil properties (Birkeland, 1999).

The influence of parent-material texture on soil development is especially well illustrated by the contrasting sequences of calcrete (i.e., calcic horizons) development in coarse- versus fine-grained parent materials in arid/semiarid environments (Gile et al., 1966; 1981; Machette, 1985; Dixon, 2013) (Fig. 5.7). Here the influence of contrasting patterns of water movement is particularly well seen (Holliday, 1994).

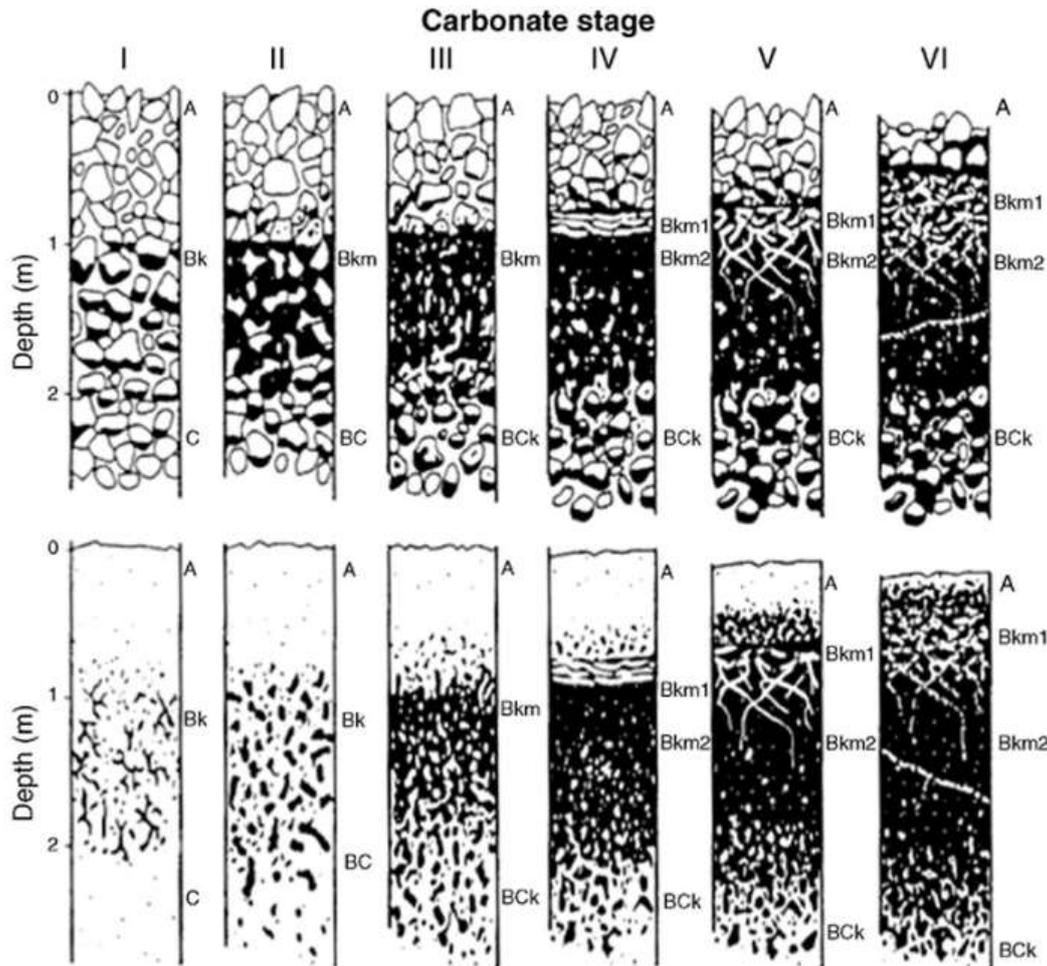


Figure 5.7 Morphological sequence of calcrete development in coarse-grained parent materials (top) and fine-grained parent materials (bottom).

The influence of contrasting parent-material mineralogies has recently been well illustrated by a study of clay minerals formed in spodosols in the Italian Alps (Mirabella et al., 2002). Here all soil-forming factors except parent material are constant. Four spodosol profiles were investigated: two on till derived from granodiorite, one on slate-dominated till, and one on tonalite till. The profiles developed on the granodiorite-derived till displayed the presence of smectite derived from chlorite and interstratified vermiculite/mica in the “E” horizon and hydroxyl-interlayered vermiculite and smectite at depth. Soils developed on slate and tonalite-dominated tills lacked the “E” horizon smectite and vermiculite and exhibited hydroxyl-interlayered vermiculite/smectite through the entire depth of the profile.

Regionally, the soil order Andisols directly reflects the influence of their parent material volcanic ash. These soils develop dominantly on volcanic glass, a parent material possessing unique physical, chemical and mineralogical properties that can change with successive eruptive events (Dixon, 2002). Upon weathering, this parent material produces products that are predominantly noncrystalline aluminosilicates known as allophane, together with opaline silica and ferrihydrite. When humus is present, allophane–humus complexes are formed. Another secondary product commonly produced from the weathering of the volcanic ash is imogolite, which together with allophane tends to dominate subsoil horizons (Buol et al., 1989).

Surface “A” horizons commonly display little evidence of weathering because of their high porosity. Under favorable drainage conditions, the subsurface horizons may develop into spodic horizons. Overall, however, these soils are weakly developed, are very freely drained, and possess low bulk densities; all direct reflections of the parent-material characteristics ( Dixon, 2002 ).

### **5.5. Soil properties and vegetation**

Of all of the state factors identified by Jenny (1941) governing the formation of soils, general agreement appears to be that the role of the biotic (i.e., vegetation) factor is the most difficult to isolate (Birkeland, 1984; Schaetzl and Anderson, 2005). Jenny regarded the biotic factor to be dependent on other state factors and to principally control biochemical processes in soils, especially those related to nitrogen ( Jenny, 1980 ). Initially the biotic factor was largely restricted to the role of plants with little consideration to the role of fauna (Schaetzl and Anderson, 2005). Vegetation patterns, although responding to climate, as well as to slope steepness and aspect, exert profound influences on moisture regulation, as well as influencing soil chemistry and mineralogy. Plants function as geochemical pumps in the soil environment. They uptake biologically necessary nutrients from the soil in solution, use them in physiological processes, temporarily store them in plant tissue, and ultimately return them to the soil in litterfall, root decay, and decomposition of the plant upon its death (Amundson et al., 2007). Vegetation affects the availability and physical distribution of nutrients in unique ways. The cycling of nutrients, such as N, P, K, and Si by plants lead to upward migration within the soil and into the vegetation, commonly resulting in the surface enrichment of these elements in the soil surface (Lucas, 2001; Jobbágy and Jackson, 2004; Amundson et al., 2007). Plant-induced ascension of silicon has been proposed as the mechanism for the formation of kaolinite-rich horizons over Fe/Al-rich horizons in the warm tropics (Lucas, 2001). The release of Si from litterfall has also been proposed as the reason for the formation of Si-rich soil surface horizons (Amundson et al., 2007). Similarly, the formation of calcrete and its associated accumulation of Ca has been widely attributed to Ca pumping by vegetation (Klappa, 1979; Kosir, 2004). Plants also play a crucial role in weathering in the soil environment largely as a consequence of the huge amounts of CO<sub>2</sub> they produce and the effects of this on water-mineral reactions and release of chemical elements. In addition, plants produce chelates, which form soluble metal complexes that facilitate increased mineral solubility. Weathering in the Critical Zone is discussed in Chapter 4 by Pope.

At the regional scale, the role of vegetation in soil genesis is strongly expressed across the boundaries between the Great Plains/Eastern Temperate forests, and the Great Plains/North American Desert Ecoregion boundaries (United States Environmental Protection Agency, 2006). The Great Plains are dominated by mollisols with the influence of a predominantly grassland vegetation strongly imprinted on soil properties (Buol et al., 1989; Dixon, 200 ). Mollisol formation is dominated by melanization, the darkening of soil by predominantly biological processes. These processes include the accumulation of organic materials from living and decaying grasses, soil mixing by organisms, eluviation and illuviation of organic materials and accompanying formation of dark-colored cutans, and the formation of dark-colored lignoproteins (Buol et al., 1989; Dixon, 2002).

Soils of the Eastern Temperate Forests are predominantly Alfisols and Ultisols (Brady and Weil, 2002). These soils are both characterized by the accumulation of clay in subsurface horizons, as well as the accumulation of organics and sesquioxides

under favorable climatic regimes (Dixon, 2002). Decay of forest litter produces ligands, which play a crucial role in the transportation of sesquioxides by chelation (Buol et al., 1989). In addition, both of these soils experience considerable profile mixing as a result of treethrow.

The Western Desert Ecoregion is characterized by the widespread development of Aridisols. These soils are characterized, in part, by the accumulation of soluble salts and the notable absence of an organic “O” horizon (Southard, 2000). The source of salts, particularly calcium carbonate, is now widely ascribed to redistribution of atmospherically derived carbonate (Dixon and McLaren, 2009). However, plants play a crucial role in the concentration of carbonate within soil profiles. The occurrence of carbonate in the vicinity of plant roots is widely documented (Klappa, 1980; Phillips et al., 1987; McLaren, 1995; Dixon and McLaren, 2009). Mumm and Reith (2007) have demonstrated the complex role played by microorganisms and plants in biomediation of calcrete as biological processes modify parent-material chemistry.

Soil/vegetation relationships are perhaps most strongly demonstrated along topographic gradients. Alexander et al. (1973) investigated the relationship between soils and vegetation along a transect from a hot oak-grassland to a very cold coniferous forest (Fig. 5.8). The accompanying soils change from Alfisols under oak-grassland and warm coniferous forest, to Ultisols under cool coniferous forest and cold coniferous forest, to Andisols under very cold coniferous forest. Accompanying these soil changes is a progressive increase in organic-carbon content.

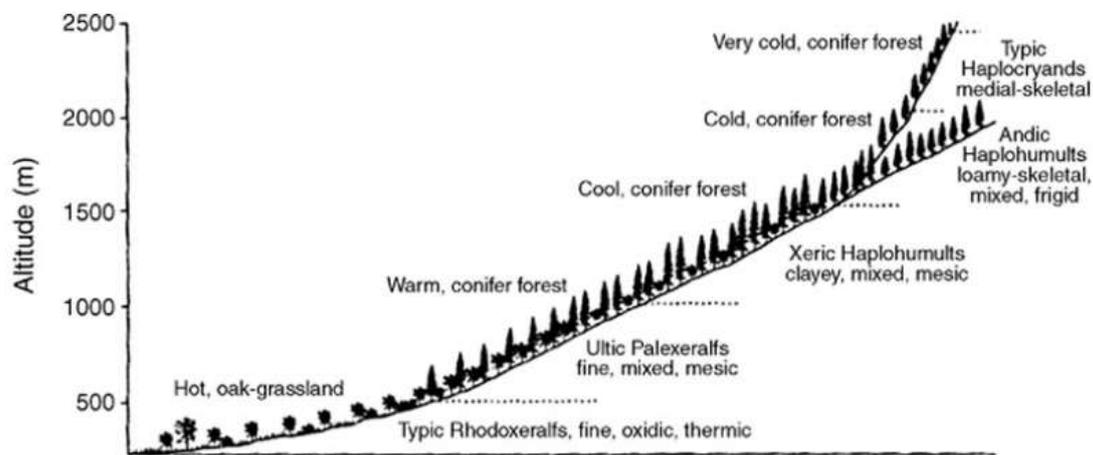


Figure 5.8 Relationship between soils and vegetation with increasing altitude in the Cascade Range of northern California.

The relationship between vegetation distribution and soil properties has recently been investigated in the mountains of northern Sweden by Darmody et al. (2004). These investigators identified strong soil/vegetation associations largely influenced by soil moisture and nutrient availability, with strong elevation/climate influences. Spodosols dominate at lowest elevations and are strongly associated with birch and Empetrum heath vegetation. Between 600 m and 900 m asl, the dominant soils are Entisols and are associated with wet meadow and willow vegetation. These soils are characterized by high CEC, pH, Ca, Mg, K, and organic C and N. Soils developed under *Dryas octopetala* (650–900 m asl) display the highest pH, base saturation and Ca contents of any of the soils in the area. The soils are predominantly

Entisols and Inceptisols, but at one location a Mollisol occurs. Soils in the alpine regions are the least well developed. Plant cover is minimal with a thin covering of cryptogams, at best. These soils display the lowest pH, base saturation, Ca, N and most of the other soil-fertility measures. The Al content of these soils is exceptionally high. Ca/Al ratios are exceptionally low, which may indicate Al phytotoxicity. The associated soils are dominantly Entisols, but where permafrost is present, gelsols are developed.

On a somewhat smaller scale, Dixon (1983) investigated the relationships between vegetation and soil chemistry and clay mineralogy of soils developed on Holocene-age soils in the Front Range of the Rocky Mountains. He found that under tundra- meadow- vegetation covers, clay mineralogy was dominated by hydrobiotite and vermiculite. Under higher life forms, such as spruce and willow, secondary minerals were dominated by smectite and interlayered smectite-vermiculite. Chemical trends in soil "B" horizons similarly reflected greater leaching under spruce and willow and less intense leaching under tundra meadow.

### **5.5.1. Pedogenesis and Time**

Soil and regolith profile development and associated complexity are strongly time dependent ( Birkeland, 1999 ; Anderson et al., 2007 ; Brantley et al., 2007 ). Birkeland (1999) pointed out that the time necessary for the production specific soil and weathering features varies. Generally speaking, however, soil properties associated with organic matter accumulation proceed rapidly whereas those associated with weathering generally proceed more slowly. Soil and regolith profiles generally thicken over time and horizon complexity increases. Increasing soil complexity is the result of progressive accumulation of materials resulting from additions, transformations and translocations as discussed previously. With time generally occurs a progressive buildup of soil profile form and constituents ( Johnson and Watson-Stegner, 1987 ). This is referred to as progressive pedogenesis. Alternatively pedogenesis may be regressive: a situation in which soils no longer change or evolve in the same direction. Curves for the buildup of soil properties are initially typically steep but eventually the curves flatten out indicating less visible change with time. This flattening out is referred to soil steady state, which may be only transitory as environmental conditions change.

### **5.6. Conclusions**

Soil constitutes a significant component of the Critical Zone. It represents the "membrane" across and through which water and solutes, energy, gases, solids, and organisms are exchanged with the atmosphere, biosphere, hydrosphere, and lithosphere. Water is the driving force of the majority of these exchanges within the soil, as well as the Critical Zone as a whole. Effective water availability is strongly controlled by climate at all scales, by topographic form and aspect, and by vegetation type and abundance. Soil formation within the Critical Zone is a function of interactions between climate, geology, and vegetation. Whereas climate and geology are largely independent state factors, vegetation-controlled processes appear to be most appropriately regarded as dependent state factors. Whereas vegetation patterns and processes respond strongly to climatic and geologic processes, they clearly have profound influences on climate at multiple scales, as well as on geologic processes.

## Chapter 6

# Soil Geochemistry in the Critical Zone: Influence on Atmosphere, Surface- and Groundwater Composition

Julia Perdrial \* Aaron Thompson \*\* Jon Chorover †

\* Department of Geology, University of Vermont, Burlington, Vermont, USA

\*\* Department of Crop and Soil Science, University of Georgia, Athens, Georgia, USA

† Department of Soil Water and Environmental Science, University of Arizona, Tucson, Arizona, USA

### Abstract

In this chapter we introduce soil biogeochemical processes that have broad influence on Critical Zone function. Soil comprises all unconsolidated material above competent bedrock that is open to – and structured by – fluxes of energy and matter. These fluxes drive biotic and abiotic reactions that result from the open system disequilibria. While (bio)geochemical reactions occur at the molecular scale, they have implications for ecosystem function and human sustainability. Many of the characteristic changes of the Anthropocene – including climatic and land use changes – have direct influence on soil biogeochemistry with cascading effects on the capacity of the Critical Zone to sustain human life. Here we highlight several of these “scalable” soil biogeochemical reactions, give examples on how they are impacted by external forcings, and show how they “reach” beyond the extent of soils to impact atmosphere, surface- and groundwater composition.

Keywords soil biogeochemistry open system molecular to global scale dissipative products climate Anthropocene

### 6.1. Introduction

The capacity of the Critical Zone to support life derives, in large part, from the structured fabric of soil. This living porous medium is formed from chemical reactions occurring among minerals, microbes, natural organic matter (NOM), solutes, and mobile fluids that are the focus of the field of soil biogeochemistry. Soil, as defined here, is thus inclusive of all unconsolidated material above bedrock (e.g., the regolith profile including mobile colluvium and saprolite). This definition of soil is consistent with early notions of soil as “a thin layer over the Earth’s surface” or “a medium for plant growth” ( Bockheim et al., 2005 ; Churchman, 2010 ; Simonson, 1968 ) and its conceptualization as “a natural body independent of surface rocks and biota” ( Coffey, 1909 ; Dokuchajev, 1883 ; Hilgard, 1882 ; Jenny, 1941 ; Shaler, 1892 ), as well as more recent concepts of soil as a landscape-scale expression of geomorphologic forces ( Daniels and Hammer, 1992 ; Huggett, 1975 ).

This broad conceptualization of soil provides a framework for scaling biogeochemical approaches and fosters development of transdisciplinary research ( Fig. 6.1 ). Soil is “the most complex biomaterial on the planet” ( Young and Crawford, 2004 ), and so it is best understood by combining multiple disciplinary perspectives. In the landscape, soil operates as a biogeochemical reactor ( Fig. 6.1 ), as discussed originally by Targulian and Sokolova (1996) . This reactor is open to matter and energy fluxes, and subject to the constraints of elemental mass-balance ( Chadwick et al., 1990 ; Crompton, 1960 ; Marshall and Haseman, 1942 ), and energy-

driven ( Minasny et al., 2008 ; Rasmussen, 2011 ; Rasmussen et al., 2005 ; Rasmussen and Tabor, 2007 ; Runge, 1973 ; Schaetzl and Schwenner, 2006 ; Volobuyev, 1964 ) concepts of pedogenesis. The soil reactor is fed by material inputs (in gas, liquid, and solid form) from weathering rock and the atmosphere, including water and carbon influxes (see Section 6.2 ). Energetic inputs include heat and reduced carbon compounds fixed through photosynthesis. Reactants are reprocessed (see Section 6.3 ) within the soil by abiotic and biologically catalyzed reactions, with impacts on surface water, groundwater and atmospheric systems resulting from their receipt of the dissipative products of these reactions. Such dissipative products include those of, for example, respiration ( $\text{CO}_2$ ), chemical and physical denudation (dissolved elements and mobilized particulates) and latent heat transfer (water vapor), that are transported through and beyond the CZ by volatilization, erosion, or water (stream/groundwater) fluxes ( Banwart et al., 2011 ; Rasmussen et al., 2011 ).

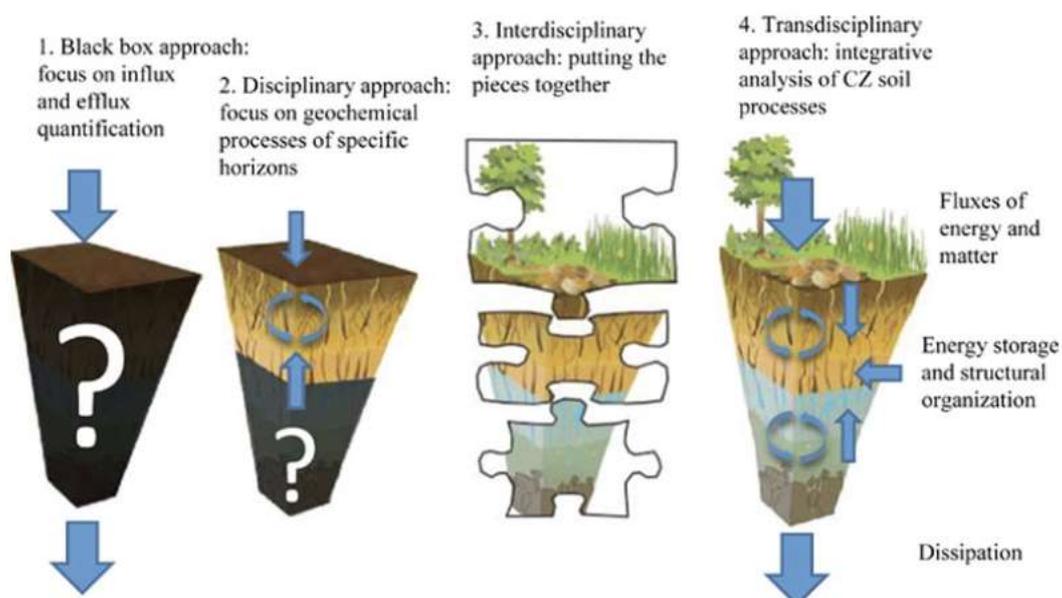


Figure 6.1 Different approaches used to study soil geochemical processes. The goal of integrative analysis of CZ processes is to combine all previous approaches in a synergistic effort

Biogeochemical processes in the soil reactor occur at the molecular scale but influence (or reach) macroscopic to global scales. For example, soils constitute the largest terrestrial carbon (C) sink and accommodate roughly three times more C than the atmosphere ( Houghton, 2007 ; Lal, 2005 ; Trumbore, 2006 ), but molecular-scale mechanisms of organo-mineral and organo-metal interaction are responsible for this C sequestration capacity (see Section 6.2.1 ). Another example is the soil's capacity for removing excess nutrients and contaminants from water, which is successfully relied on for on-site wastewater systems (i.e., septic tank leachfields, USEPA (2002) ) and retardation of radioactive waste ( Yasunari et al., 2011 ; Zachara et al., 2007 ). A suite of molecular-scale processes including sorption and transformation reactions (see Section 6.3 ) are responsible for providing this important service of “natural attenuation”. Finally, it is molecular scale soil processes that provide ecosystems with life sustaining nutrients, including the fixation of nitrogen via microbially mediated redox reactions ( Sections 6.2 and 6.3.3 ), and the dissolution of P and other lithogenic

macro- and micronutrients (e.g., Ca, Mg, K, Fe, Zn, etc.) via mineral weathering reactions (see Section 6.3.4 ). Chemical weathering in soils is the ultimate sink for atmospheric CO<sub>2</sub> in the geological carbon cycle, as it results in proton consumption that enables carbonate release to rivers and oceans where it is subsequently buried in marine sediments, a process of critical importance to climate sustainability ( Gaillardet et al., 2011 ; Gislason et al., 2009 ).

## 6.2. Material for soil geochemical reactions

Geochemical elements are fed to the soil reactor through multiple pathways that include biologic fixation, wet and dry deposition, and bedrock weathering. In this section we describe the processes leading to the introduction and stabilization of C, N and lithogenic elements. Many of the elements required for life derive from rock weathering (Ca, Mg, K, Fe, Zn, etc.), whereas others are present only at trace amounts in silicate rock (e.g., C, N) and therefore derive from biological fixation of atmospheric species ( Fig. 6.2 ).

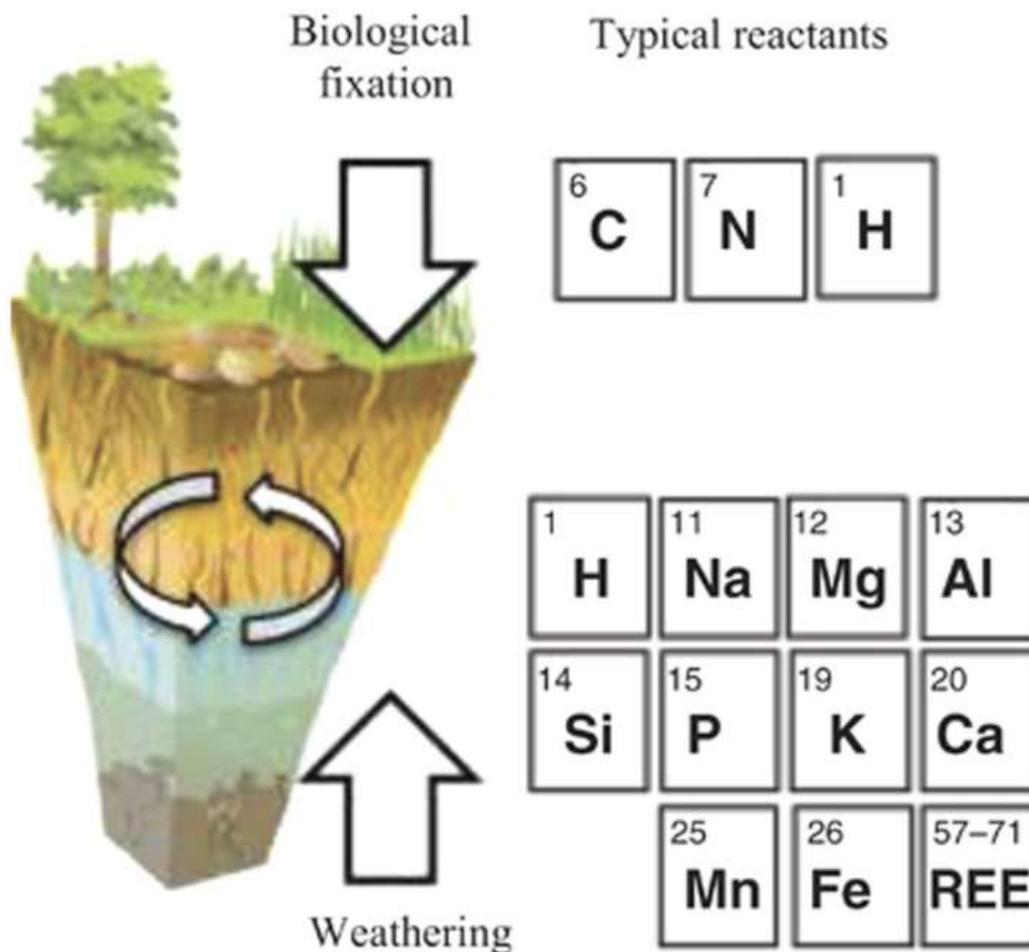


Figure 6.2: Sources of material for soil biogeochemical processes

### 6.2.1. Carbon

Natural Organic Matter (NOM) enters the CZ via photosynthesis by autotrophic organisms (dominantly vascular plants) that transform atmospheric CO<sub>2</sub> into biomass and thereby provide a constant supply of biomolecular material in the form of decaying plant and microbial debris ( Essington, 2002 ; Sparks, 2005 ). As a

result, soils are significantly enriched in C relative to parent rock ( Sposito, 2008 ). As a reduced substrate, NOM is the primary source of energy and carbon for heterotrophic microorganisms, and the reactions fueled by NOM inputs underpin numerous foundational soil biogeochemical reactions as discussed in Section 6.3 . Heterotrophic soil microbes oxidize this reduced C, producing soil partial pressures of CO<sub>2</sub> (PCO<sub>2</sub>) that are highly variable in space and time, and that can exceed atmospheric values by orders of magnitude with important implications for soil geochemistry (see Section 6.3 ). Different pools of C can be distinguished based on their origin and persistence. For example, plant debris, leaf litter, and animal carcasses are collectively termed organic residues and form one important pool of young, soil organic materials ( Essington (2002) , Fig. 6.3 ). A wealth of living organisms depend on this pool including macro and meso fauna as well as microorganisms, which themselves represent a second pool of organic material, termed soil biomass ( Sparks, 2005 ). The third pool, broadly termed soil organic matter (SOM), is a result of the decomposition and rearrangement of materials from the first two pools and is more persistent in soils.

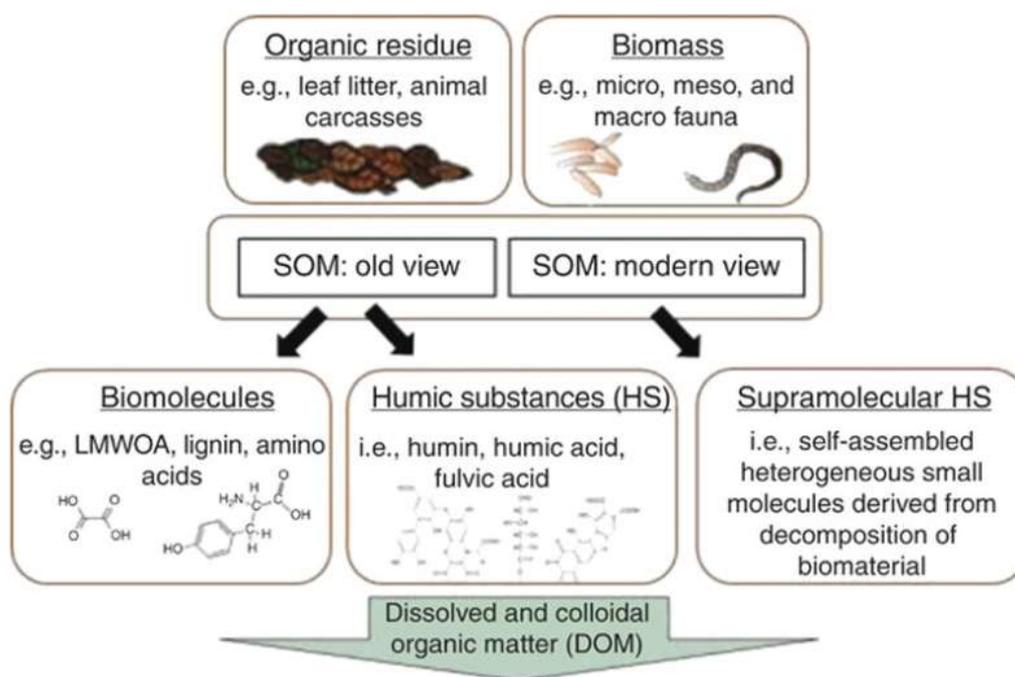


Figure 6.3 Schematic sketch of SOM showing the old view and new view of HS.

Traditionally, SOM was separated into two groups: SOM components that could be recognized as biochemical molecular classes were termed biomolecules (or nonhumic substances); whereas molecules with only remote resemblance of their precursor material were defined as humic substances (HS). The latter were long considered to represent high molar mass (>10,000 Da) polymerized degradation products of biomolecules. However, understanding of the molecular structure of SOM has advanced and more recent research suggests that SOM consists of smaller molecules, dominantly biomolecular fragments, that are aggregated to supramolecular structures ( Kelleher and Simpson, 2006 ; Piccolo, 2002 ). In these structures, molecules are bound together via hydrophobic interaction, hydrogen bonding, and

cation bridging that impart macromolecular characteristics of high molecular weight compounds ( Chorover, 2011 ; Leenheer et al., 2001 ; Sutton and Sposito, 2005 ).

The modern view of SOM is central to understanding soil C dynamics since all previously mentioned bonding mechanisms are based on relatively weak interactions, rendering these SOM aggregates highly reactive and dynamic. Solution geochemistry exerts a strong control on these supramolecular entities where, for example, the presence of polyvalent cations increases the tendency for aggregation whereas the presence of low molecular weight organic acids (LMWOAs) has the opposite effect ( Piccolo, 2002 ; Simpson, 2002 ).

Dissolved organic matter (DOM) represents that fraction of the waterborne organic matter (OM) pool that is, in most cases, defined operationally as passing through ca. 0.45  $\mu\text{m}$  filtration. Although DOM represents only a small fraction of total SOM, it is the most mobile and influences most terrestrial and aquatic biogeochemical processes ( Bolan et al., 2011 ; Gabor et al., 2014 ; McGill et al., 1986 ). The characteristics of DOM are related to their source (e.g., throughfall, leaf leachate, soil pore water, hyporheic zones, and groundwater, see Inamdar et al. (2012) and Singh et al. (2014) ). For each of these source reservoirs, molecular spectroscopic indices has shown systematic variation with location in the CZ. For example, the fluorescence index, a measure of OM provenance ( McKnight et al., 2001 ), indicates microbially-derived DOM increases as a fraction of total DOM for waters exhibiting longer residence time in the CZ. Conversely, UV specific absorbance (i.e., absorbance at 254 nm normalized to dissolved organic C concentration), an indicator for DOM aromaticity ( Weishaar et al., 2003 ), decreases systematically from soils to GW ( Inamdar et al., 2011 ).

NOM composition changes during its journey through the CZ as different sources of DOM mix and they are acted on by microbial cycling and abiotic fractionation processes (processes that preferentially remove certain portions of DOM). The composition and character of NOM products in stream and groundwater are therefore source and process dependent ( Boyer et al., 1997 ; Perdrial et al., 2014 ; Raymond and Saiers, 2010 ; van Verseveld et al., 2008 ). The fractionation of DOM by mineral surfaces, for example, strongly influences mobile OM composition. The relative retention of higher molar mass, aromatic fractions of DOM on soil particle surfaces has been investigated experimentally using metal (oxyhydr)oxides, smectite, and whole soils ( Chorover and Amistadi, 2001 ; Guo and Chorover, 2003 ; Vázquez-Ortega et al., 2014 ). OM that is not attenuated is transported from soil as dissolved and colloidal products to influence surface- and groundwater composition. As mentioned C plays multiple roles in soils and is involved in many different soil biogeochemical reactions that are detailed in Section 6.3 .

Biomolecules delivered to soil through autotrophic synthesis and microbial turnover exhibit varying degrees of lability to decomposition. When free from mineral and metal complexes, certain classes of biomolecules, such as lignin, exhibit greater kinetic stability that slows (but does not eliminate) microbial degradation relative to other classes of biomolecules, such as saccharides that are more readily mineralized to  $\text{CO}_2$  . Degradation rates of SOM, therefore, exhibit strong dependence on the relative concentrations of various biomolecular types ( Berg and Meentemeyer, 2002 ; Talbot et al., 2012 ). However, as suggested earlier and discussed in more detail later, SOM can be protected from microbial processing by its interactions with other co-aggregated soil components and, therefore, even inherently labile biomolecular forms can be stabilized over long time scales ( Houghton et al., 1998 ; Kaiser and Guggenberger, 2000 ; Kleber and Johnson, 2010 ; Lal, 2004 ; Schmidt et al., 2011 ).

In both cases, molecular-scale processes are ultimately responsible for the formation of a global-scale C sink.

### 6.2.2. Nitrogen

Like carbon, nitrogen (N) is also greatly enriched in soils relative to parent rock ( Sposito, 2008 ). This is because although it is abundant as  $N_2$  gas in the Earth's atmosphere (78%), oxidizing it to soluble forms (e.g. N-fixation) requires enzymatic catalysis. Enzymatic catalysis of N-fixation can be accomplished by either free-living, N-fixing bacteria such as *Azotobacter* or by bacteria living in close symbiosis with roots of leguminous higher plants, such as rhizobia . N-fixation also can occur abiotically by lightning, but this pathway is small relative to those mediated microbially. Because the fixation process comes with a high metabolic cost, N-fixation is closely linked to the aerobic respiration of organic C ( Schlesinger and Bernhardt, 2013 ).

The dominant soil N species and its mobility depend largely on environmental conditions. For example under conditions of low redox potential (  $E_h$  ) that are prevalent in waterlogged soils, the reduced N-species,  $NH_4^+$  , dominates and, as a cation, can interact via adsorption (see Section 6.3 ) with negatively charged exchange sites on clay minerals or OM that predominate in temperate zone soils. Under oxic conditions, aerobic nitrification reactions mediated by nitrifying bacteria oxidize  $NH_4^+$  to produce nitrate ( $NO_3^-$  ) ( Barsdate and Alexander, 1975 ; Kowalchuk and Stephen, 2001 ). As an anion,  $NO_3^-$  is subject to electrostatic repulsion from the same types of negatively charged surface sites that retain  $NH_4^+$  and, as a result, it is readily transported through soil profiles that comprise dominantly negatively charged soil particles. Conversely,  $NO_3^-$  is adsorbed and  $NH_4^+$  is mobilized in soils comprising predominantly positively charged surface sites, as is the case for highly weathered tropical soils (e.g., oxisols) that are dominated by Fe and Al (oxyhydr)oxides and kaolin group minerals. Organic N in soils is mostly present as peptide (amide) N and is, as an important component of SOM, closely linked to the fate of SOM and the reactive transport of DOM more generally. Although N is added to the soil profile from the atmosphere, it is recycled many times to meet CZ nutrient requirements ( McNeill and Unkovich, 2007 ). Characteristic of the Anthropocene, global N cycling reflects a large human imprint; anthropogenic activities have altered the nitrogen cycle of many CZ ecosystems ( Vitousek et al., 1997 ). To a greater extent than other elements, the global N cycle is affected by its delivery to soils as fertilizer or in the form of nitrogen oxides ( $NO_x$  ) via wet and or dry deposition downwind from industrialized areas ( Likens et al., 1972 ; Rice and Herman, 2012 ).

### 6.2.3. Lithogenic Elements

Elements derived from the weathering of rock, i.e., lithogenic elements, are transferred to the soil incipiently at the weathering front (i.e., bedrock–regolith interface), which resides at the bottom of the weathering profile ( Brantley et al., 2007 ; Holbrook et al., 2014 ). As discussed in Chapter 4 , weathering of primary minerals supplies the CZ with silicic acid and soluble metals (Group 1 and 2 cations, transition metal cations, nutrient and toxic metal(loid)s and nonmetals including P, S, Se and As and other trace metal(loid)s including the rare earth elements, etc.). Weathering consumes protons and/or  $CO_2$  and  $O_2$  in proton-promoted and oxidative weathering reactions, respectively. Rock-derived elements released to solution are either: (1) recycled in the CZ – as is the case for bio-essential lithospheric elements (e.g., Ca, K, and Si, Amundson et al. (2007) ); (2) coprecipitated during the formation of

secondary minerals (e.g., Al, Fe, and/or Si in clay minerals and/or metal (oxyhydr)oxides, (see Section 6.3 ); or (3) released to surface waters as dissipative products of CZ evolution (e.g., Na, Mg, bicarbonate).

Lithogenic element accretion in soils also results from aeolian deposition, which includes salts as well as mineral dust ( Derry and Chadwick, 2007 ). Dust deposition can be the dominant soil source of lithogenic elements in arid regions where weathering rates are low, and it can also be an essential contributor to nutrient demands of ecosystems in highly weathered soils that have become depleted in elements such as P and Ca ( Derry and Chadwick, 2007 ). Especially in nutrient poor soils, for example, intensely weathered soils in Hawaii ( Carrillo et al., 2002 ; Derry and Chadwick, 2007 ) or soils on the Colorado plateau ( Reynolds et al., 2006 ), deposition can supply the majority of nutrients. Atmospheric deposition from anthropogenic activities such as smelting operations and fossil fuel combustion contribute to the near surface accumulation of metals such as Mn and Pb ( Herndon et al., 2010 ).

### 6.3. Soil biogeochemical reactions and their impact on the composition of atmosphere, surface- and groundwaters

#### 6.3.1. Soil Fluids Determine the Reach of Biogeochemical Processes

Soils are variably saturated, aqueous heterogeneous systems where water and air serve as the fluid media for material exchange that fuels microbial activity, mineral weathering, nutrient and toxic element transformation, organic compound biodegradation, and their composite effects on soil formation. Both gaseous and aqueous phases of soil contain and transport reactants/products throughout different segments of the CZ and to external portions of the Earth System (e.g., atmosphere, hydrosphere; Perdrial et al. (2014) ). Soils with high water throughput become depleted of mobile elements ( Berner and Berner (2012) and references herein) ( Fig. 6.4 ). In contrast, soils with low water throughput accumulate relatively soluble solids, including for example, carbonate, sulfate or even (under extremely arid conditions) nitrate and chloride minerals along with co-precipitated cations (Ca, Mg, Na) because precipitation is small relative to potential evapotranspiration and hence weathering products are not removed. Because of its functionality as a polar and weakly acidic solute, water is itself a reactant in many important CZ reactions and processes ( Fig. 6.4 ). For example, water hydrates solids and solutes; self-dissociates/protonates as a weak acid; and serves as a reactant in several kinds of soil geochemical reactions, including water incorporation into secondary minerals, that are detailed in the following sections.

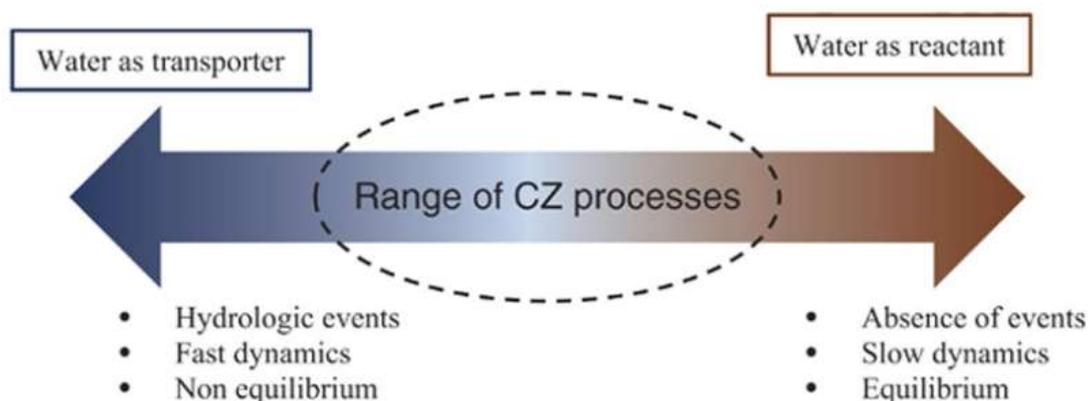


Figure 6.4 Conceptualization of two end member roles of water that impact soil geochemical processes, connects different Critical Zone (CZ) segments, and transfers external forcings to the CZ.

Like the above-ground atmosphere, the soil gas phase contains  $N_2$ ,  $O_2$ , and  $CO_2$  as well as other trace gases (Ar, NO,  $N_2O$ ,  $NH_3$ ,  $CH_4$ ,  $H_2S$ ), but their partial pressures can differ significantly over space and time, with much greater heterogeneity than that of the well-mixed, above-ground atmosphere. Globally relevant trace gases are produced and consumed by microbial activity in soils, with important feedbacks on atmospheric chemistry. Of particular importance are soil-derived “greenhouse” gases that contribute to the radiative forcing associated with global warming ( $N_2O$ ,  $CH_4$ ) that are produced when  $O_2$  partial pressures limit its availability as a terminal electron acceptor in heterotrophic microbial respiration. The influences of soil (bio)geochemical processes on the chemical composition of atmospheric and surface water reservoirs are manifold. The impact of reactions within soils on reservoirs external to the CZ is termed here the “reach” of a soil biogeochemical process, and we include examples of these in the discussion of each biogeochemical process.

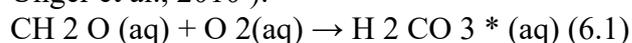
### 6.3.2. Element Speciation in Soil Pore Fluids

#### 6.3.2.1. Overview

Water resides in soil pore spaces. In the unsaturated (vadose) zone, pore spaces are subject to fluctuations between wetting and drying conditions. After wetting events, as pore spaces become progressively drier because of evapotranspiration, solute concentrations increase, leading to precipitation of soluble salts including sulfates, carbonates and, under extremely arid conditions, even nitrates or chlorides. However, during wetting front propagation, as occurs following a rainfall or snowmelt event, these salts are rapidly re-solubilized as pores become progressively filled with aqueous solution. During such events, particularly in water-limited CZ systems, the increased availability of moisture often promotes a cascade of biological activity including root growth, root exudation and border cell production. This release of labile OM into soil pores, along with the downward percolation of DOM from the litter layer, drives heterotrophic microbial metabolism. As discussed later, the effects of such dynamics drive changes in the speciation of elements dissolved in soil solution and this, in turn, affects their transport and fate throughout and beyond the CZ.

#### 6.3.2.2. Soil Geochemical Processes and Agents

The enhanced heterotrophic microbial activity that occurs in the presence of water during downward propagation of the wetting front results in the release of dissolved  $CO_2$  plus carbonic acid (together denoted  $H_2CO_3^*$ ) to soil solution because of the decomposition of SOM (i.e.,  $CH_2O$ ) (Huxman et al., 2004 ; Placella et al., 2012 ; Unger et al., 2010 ):



$H_2CO_3^*$  is also produced as a result of root respiration of  $CO_2$  followed by its dissolution (i.e., hydration) by water ( Fig. 6.5 ). This dissolved carbon dioxide plus carbonic acid ( $H_2CO_3^*$ ) is subject to two principal fates upon its formation as a result of enhanced biological activity. It can shed its hydration water and partition

out of the soil solution and into the soil gas phase where it gives rise to partial pressures of CO

2(g) that may be orders of magnitude higher than the above-ground atmosphere:



OR carbonic acid may dissociate in the soil solution to form a reactive proton that contributes to acidifying soil solution, along with a bicarbonate ion that can be discharged to ground or surface waters.

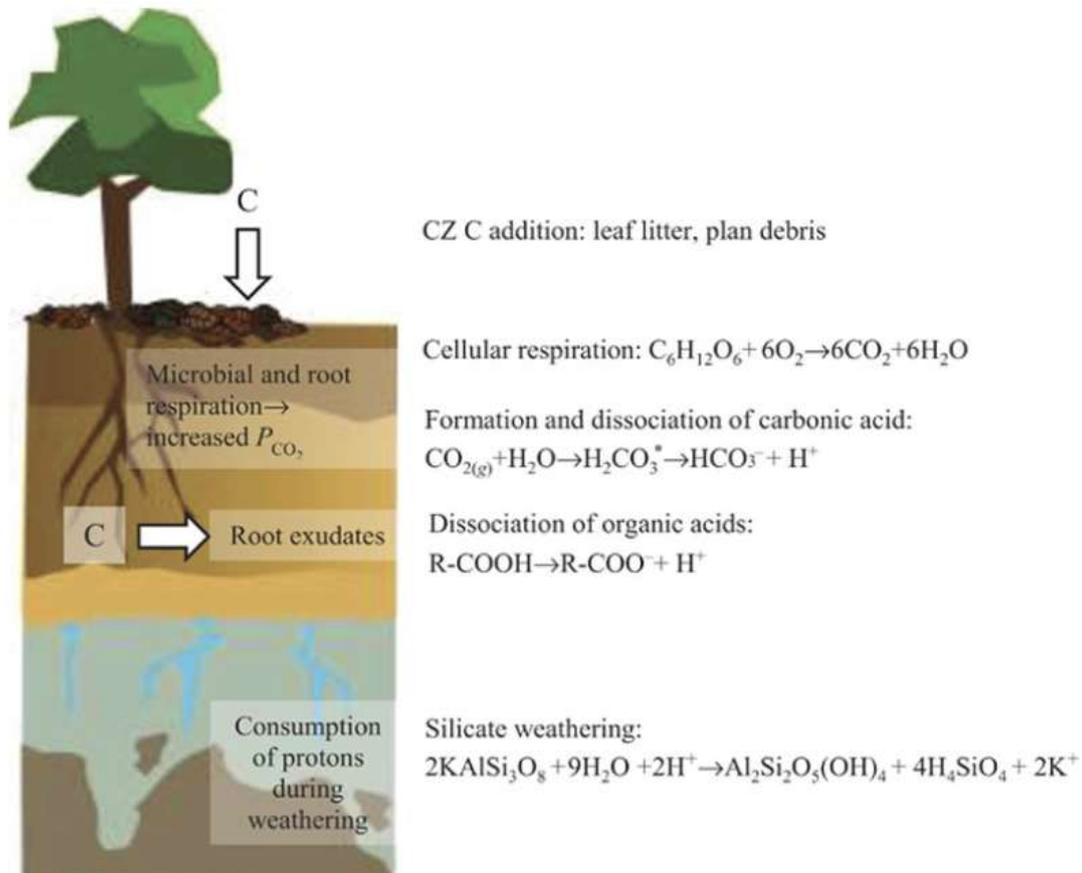


Figure 6.5 Schematic view of carbon (C) cycling in soils and its impact on the proton balance for biogeochemical reactions.

These distinct fates highlight the “reach” of soil biogeochemistry to the atmosphere or surface waters external to the CZ, because they either result in CO 2(g) return to the atmosphere Eq. (6.2) or its “consumption” via proton-promoted Eq. (6.3) weathering reactions (discussed later) that lead to the transport of bicarbonate ion out of the weathering zone and into rivers and oceans. The combinations of Eqs (6.1) and (6.2) describe the commonly observed “Birch effect” wherein soil wetting, particularly in water-limited systems, results in a pulsed release of CO 2(g) from the soil to atmosphere ( Fierer and Schimel, 2002 ).

Aqueous phase reactions that occur during pore filling are analogous to those pertaining to natural waters in general. These reactions involve water (solvent), and the dissolved solutes, which are inorganic and organic, and include cationic, anionic, and neutral molecules ( Sposito, 2008 ). Initial wetting front propagation leads to simple hydration of cations and anions by water, and this is the dissolution reaction

that brings “soluble salts” back into solution. As a dipolar solvent, water forms stable coordination complexes with cations through the electronegative oxygen and with anions through the electropositive hydrogen groups of the water molecule ( Richens, 1997 ). These “hydration complexes” of, for example,  $\text{Ca}^{2+}$  or  $\text{K}^{+}$  with water (i.e.,  $\text{Ca}(\text{H}_2\text{O})_6^{2+}(\text{aq})$ ) are one example of metal-ligand complexation, where the metal (e.g.,  $\text{Ca}^{2+}$ ) forms a coordination complex with the electronegative oxygen of the ligand ( $\text{H}_2\text{O}$ ).

Hydrated ions can undergo aqueous phase reaction with other ions in solution, including complexation and/or electron transfer ( Richens, 1997 ). The former occurs when ligands form outer-sphere (hydration shell retained) or inner-sphere (hydration shell displaced) complexes with ligands other than water. For example, hydrated  $\text{Ca}^{2+}(\text{aq})$  can form an outer-sphere complex (i.e., ion pair) with the sulfate anion to form the neutral  $\text{CaSO}_4^0(\text{aq})$  species. Such reactions are highly relevant to the role of soils in CZ geochemistry, because, for example, the formation of a neutral dissolved species such as  $\text{CaSO}_4^0(\text{aq})$  makes Ca more mobile than it would be as the free aquoion  $\text{Ca}^{2+}(\text{aq})$ , which can be adsorbed to negatively charged cation exchange sites. Hydrated polyvalent metal cations such as  $\text{Al}^{3+}(\text{aq})$  or  $\text{Fe}^{3+}(\text{aq})$  form strong (often inner-sphere) coordination complexes with dissolved organic ligands, including low molecular mass organic acids such as oxalic, phthalic or mallic acids. Again, such complexation with anionic ligands diminishes overall molecular charge on the metal–ligand complex, making the metal more mobile in the soil system, transporting it to depth, and/or to ground or surface waters. Indeed, the complexation of Al or Fe with organic ligands is a key reaction in podsolization: the process of Spodosol formation wherein metal-ligand complexes are “eluviated” from surface “E” horizons and precipitated in subsurface humic (Bh) or spodic (Bhs) horizons ( Lundstrom et al., 2000 ).

### **6.3.2.3. Reach**

Aqueous phase speciation has important implications for the fate and transport of lithogenic (e.g., Ca, Al, K, Fe) and biologically fixed atmospheric (e.g., C, N) elements in soils and throughout the CZ. Complexation reactions alter the charge, stereochemistry, and reactivity of ions in solution. The impacts are manifest in the genesis of soil profiles, such as the complexation, translocation, and aggregation of metal–ligand complexes. But they also impact more broadly the biogeochemistry of the CZ and beyond. For example, concentration-discharge relations of surface waters in many forested catchments show correlated catchment releases of polyvalent metals and DOM with characteristic stoichiometries of metal–ligand complexes ( Chorover et al., 2011 ; Gangloff et al., 2014 ). Likewise, the partitioning of  $\text{CO}_2$  between water- or gas-filled soil pores has important implications for the net exchange of carbon between land and atmosphere.

## **6.3.3. Adsorption–Desorption Reactions**

### **6.3.3.1. Overview**

Soils comprise a diversity of mineral and organic solids with reactive surface functional groups. Molecules that are dissolved in aqueous solution are termed “adsorptive” because they can potentially accumulate at these interfaces. The process of accumulation at interfaces, termed “adsorption,” results in the formation of surface-associated species that are retained against leaching, and retarded in their transport by advected water. Such adsorbed species, “adsorbate,” serve as a reservoir

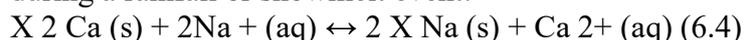
of various bioavailable nutrient, toxic or inert species subject to potential biological uptake or subsequent mobilization. The release of adsorbate back into solution (i.e., “desorption”) occurs in response to pore water perturbations that are common in the open CZ system.

### 6.3.3.2. Soil Biogeochemical Process and Agents

Soils are very complex and heterogeneous mixture of fine particulate matter comprising minerals, microbial biomass, macrofauna and root tissue, as well as SOM in various stages of bio-decay, all admixed into multicomponent aggregates. Because of the small size of solid-phase weathering products such as secondary clays, a relatively small mass of soil can contain a massive amount of surface area. For example, one gram of soil rich in the secondary aluminosilicate, smectite, can contain up to hundreds of meters of highly reactive surface available for sorption reactions.

The majority of surface area in soil is composed of (1) layer silicate clays (e.g., vermiculite, montmorillonite, kaolinite), (2) metal (oxyhydr)oxides (e.g., gibbsite, goethite, ferrihydrite, birnessite), and (3) SOM. All of these solids comprise surface functional groups, that is, small groups of atoms combined into structural units that extend from the solid surface into contact with the solution phase. The types of surface functional groups on a given particle govern a particle’s reactivity toward solution phase molecules, including nutrients and pollutants ( Sposito, 2004 ).

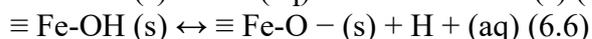
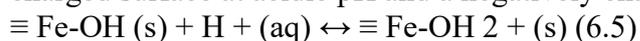
Layer silicate clay surfaces are dominantly negatively charged as a result of isomorphic substitution of lower valence ions for higher valence ions in the crystal structure. For example, Al<sup>3+</sup> for Si<sup>4+</sup> substitution is common in the tetrahedral sheets of vermiculite and beidellite whereas Mg<sup>2+</sup> for Al<sup>3+</sup> is common in the octahedral sheets of montmorillonite. This isomorphic substitution gives rise to a negative structural surface charge that manifests at planar interlayer siloxane sites of these clay minerals, and that serves as important location for cation adsorption and retention against leaching. The permanent negative charge of siloxane sites must be balanced by an equal magnitude of oppositely charged adsorbate, which can undergo cation exchange upon a perturbation of solution chemistry, for example, as occurs during a rainfall or snowmelt event:



where, X represents one mole of negative structural surface charge.

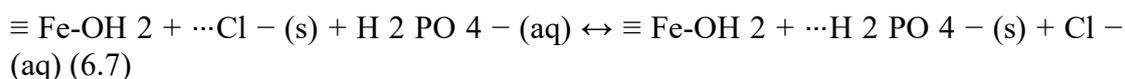
Equation (6.4) depicts a heterovalent ion exchange reaction where one mole of Ca<sup>2+</sup> adsorbate is exchanged and released to solution upon adsorption of two moles of Na<sup>+</sup>. Such cation exchange reactions in soils exert strong control over dynamic changes in pore water chemistry that occur during hydrologic events.

Whereas the metal (oxyhydr)oxides and SOM constituents do not contain significant permanent, structural charge, they have weakly acidic surface functional groups that undergo proton adsorption and desorption leading to variable surface charge as a function of soil pH. For example, the mineral goethite ( $\alpha$ -FeOOH) comprises surface hydroxyl groups with a Brönsted acidity that confers a positively charged surface at acidic pH and a negatively charged surface at alkaline pH.

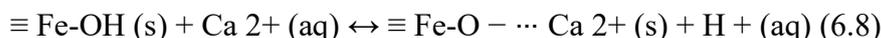


where, the “ $\equiv$ ” denotes the surface where the surface functional group (e.g., FeOH). It is noteworthy that the charge on such particles can be dynamic in the open CZ system because of the change in proton activity that may accompany respiration and carbonic acid production and dissociation Eq. (6.3). Particle surface charge then exerts strong

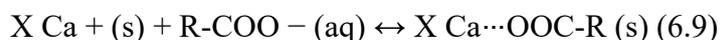
control over the adsorption–desorption of charged ions in solution. For example, goethite can serve as an effective adsorbent for anions at low pH (<8):



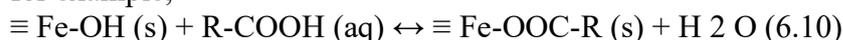
where, “…” denotes surface bond formation. In this case, the  $\text{Cl}^-$  adsorbate is subjected to desorption and displacement by  $\text{H}_2\text{PO}_4^-$ . Such Fe (oxyhydr)oxides can serve as effective adsorbents for cations at high pH (>8):



In addition to adsorption–desorption reactions of inorganic ions, both layer silicate clays and (oxyhydr)oxides are important adsorbents for SOM. These SOM adsorption reactions are central to the capacity of soils to retain the largest mass fraction of organic carbon at the Earth’s surface. Adsorbed (exchangeable) cations, such as those depicted in Eq. (6.4), play a central role in SOM adsorption to layer silicate clays by bridging between the negatively charged siloxane sites of the mineral surface and the negatively charged sites (e.g., carboxyl,  $\text{R-COO}^-$ ) of SOM:



Adsorption to Fe and Al (oxyhydr)oxide surfaces is considered to be of particular importance for the retention of SOM. This is attributed to ligand exchange reactions at hydroxylated surfaces, where surface hydroxyl groups are exchanged with carboxyl groups of SOM enabling inner-sphere complexation of SOM with metals, for example,



Such adsorbate–adsorbent bonds formed between pedogenic clays and organic ligands are critically important to the stabilization of OM against degradation in soils.

Similar reactions to those shown in Eqs (6.6) and (6.8) occur on the surfaces of solid phase SOM since SOM comprises dominantly carboxylic ( $\text{R-COOH}$ ) and phenolic (aromatic-OH) acid functional groups that likewise exhibit pH-dependent (variable) charge. In addition to its role in cation adsorption and exchange, OM is the principal sink for hydrophobic organic compounds (HOCs) introduced to soils, including many pollutants like polycyclic aromatic hydrocarbons deriving from fossil fuel combustion. These HOCs are adsorbed to OM via van der Waals interactions with uncharged moieties (aromatic rings and aliphatic chains) of the OM (Chorover and Brusseau, 2008).

### 6.3.3.3. Reach

Surface reactions in soils control the bioavailable pool of ions for plant and microbial uptake and represent a store of nutrients for vegetative growth. The composition of the exchange complex (e.g., nutrient cations such as K, Ca, and Mg versus toxic Al) feeds back to affect nutrient status and vegetation type, whereas adsorbate exchange reactions (e.g., Eq. 6.4) also control concentration-discharge relations observed in effluent surface waters. Adsorption of OM to clay particle surfaces leads to formation of heteroaggregates and the high spatial variation in geochemical conditions in soils (complex pore size distributions, steep biogeochemical gradients or proximal oxic and anoxic domains).

### 6.3.4. Reduction–Oxidation Reactions

#### 6.3.4.1. Overview

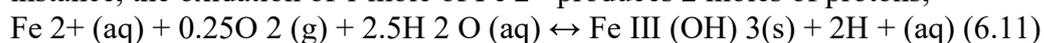
Wetting front propagation into the unsaturated portion of the CZ, as occurs during rainfall or snowmelt events, leads to pore filling with water. Consumption of dissolved oxygen results from heterotrophic respiration Eq. (6.1) such that if bio-active soils remain saturated for prolonged periods, they become suboxic and a cascade of alternative (to  $\text{O}_2$  reduction) terminal electron accepting processes (TEAPs) ensues (Borch et al.,

2010 ). However, a diversity of TEAPs frequently occurs even in soils that have bulk oxic character because of their structural complexity and heteroaggregate structure ( Blagodatsky and Smith, 2012 ; Renault and Stengel, 1994 ). Such soils comprise microsites with oxic, suboxic, and anoxic character often in close proximity (within millimeters) of each other. Steep biogeochemical gradients produce a thermodynamic driving force for reactions that would be mutually exclusive in well mixed systems. As a result, soils can host in close proximity a diverse array of microbes that use different TEAPs for respiration.

#### 6.3.4.2. Soil Biogeochemical Process and Agents

Rainfall incident on the vegetated CZ infuses water and labile OM into microbially active soils where the oxidation of OM promotes the consumption of gaseous and dissolved O<sub>2</sub> in respiration Eq. (6.1) . When molecular oxygen becomes limiting to respiration, facultative anaerobic microbes turn to anaerobic respiration, which is a form of respiration using electron acceptors other than oxygen. Whereas molecular oxygen is a strong oxidant, anaerobes use other weaker oxidants such as NO<sub>3</sub><sup>-</sup>, Mn(IV), Fe(III), SO<sub>4</sub><sup>2-</sup>, CH<sub>2</sub>O, and CO<sub>2</sub>. These terminal electron acceptors have progressively smaller reduction potentials than O<sub>2</sub>, meaning that progressively less energy is released per mole of oxidized species. Anaerobic respiration is, therefore, in general energetically less efficient than aerobic respiration. However, the processes have important implications for soil geochemistry in the CZ.

Possible redox reactions are dictated by the state of the soil system and the corresponding Gibbs energies of reaction. Among the reactions that are possible thermodynamically, those that predominate at any given point in time are determined by redox kinetics, the latter being governed to a large degree by microbial catalysis. A sequence of TEAPs is observed typically along redox gradients both spatially and temporally in soil systems. This sequence corresponds closely to progressive decreases in the Gibbs energy of the full redox reaction. Since O<sub>2</sub> is sparingly soluble in water (0.25 mM at 25°C and atmospheric partial pressure) it can be depleted rapidly by microbial and root respiration in soils subjected to limited influx of air or oxygenated water. When that occurs, dissolved nitrate and available Mn(IV) solids are utilized as alternative electron acceptors during oxidation of OM (termed nitrate reduction or denitrification and Mn reduction, respectively). As these reactants become depleted, further reduction in redox potential ( E<sub>h</sub> ) results in the successive use of Fe(III) solids (ferric reduction), SO<sub>4</sub><sup>2-</sup> (sulfate reduction), and eventually OM itself (fermentation or methane fermentation) or CO<sub>2</sub> (methanogenesis). In each case, depletion of reactant oxidizing agents and accumulation of their reduced-form products (e.g., Mn<sup>2+</sup>, Fe<sup>2+</sup>, HS<sup>-</sup>, CH<sub>3</sub>COO<sup>-</sup>, CH<sub>4</sub>) diminishes the energy yield of a given TEAP as the Gibbs energy for the full reaction approaches zero ( Chadwick and Chorover, 2001 ). Importantly, redox reactions involving metals can bring about large shifts in pH as the metals undergo hydrolysis reactions. For instance, the oxidation of 1 mole of Fe<sup>2+</sup> produces 2 moles of protons,



thus, impacting adsorption-desorption reactions, microbial activity, and other pH dependent processes, including carbon availability ( Hall and Silver, 2013 ). The reverse reaction leads to an increase in pH ( Kogel-Knabner et al., 2010 ) and has been implicated in mobilizing OM ( Buettner et al., 2014 ; Grybos et al., 2009 ; Pédrot et al., 2009 ), nutrients ( Henderson et al., 2012 ), and metals ( Davranche et al., 2011 ; Weber et al., 2009 ) as pH-driven changes in particle surface charge promote colloid dispersion ( Ryan and Gschwend, 1994 ).

#### **6.3.4.3. Reach**

Production of CO<sub>2</sub> by a variety of TEAPs results in equilibrium partial pressures of CO<sub>2</sub> in soils that exceed that of the above-ground atmosphere by orders of magnitude. This drives mineral weathering reactions (see Section 6.3.5) and the associated long-term atmospheric CO<sub>2</sub> drawdown. While silicate weathering consumes CO<sub>2</sub>, a potent greenhouse gas, N and C based TEAPs can also significantly promote radiative forcing. Methanogenesis, as occurs in soils subjected to prolonged anoxia, generates a greenhouse gas that is 30 times more effective than CO<sub>2</sub> at trapping heat in the atmosphere. Likewise, denitrification can produce gaseous intermediates that include N<sub>2</sub>O, a greenhouse gas 300 times more effective than CO<sub>2</sub> at trapping heat in the atmosphere. Reduced forms of Fe<sup>2+</sup> and Mn<sup>2+</sup> are more soluble than their oxidized counterparts, and their removal from the soil by leaching results in a permanent loss in these lithogenic ions and their oxidizing capacity.

### **6.3.5. Mineral Weathering Reactions**

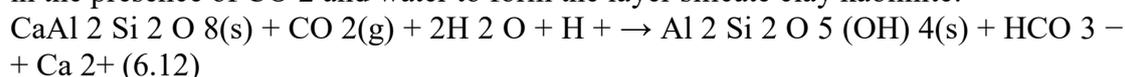
#### **6.3.5.1. Overview**

The genesis of soil involves biogeochemical weathering (i.e., formation of secondary minerals from primary minerals) in addition to the physical disintegration of rock (e.g., fractures along grain boundaries, freeze-thaw, etc.). The focus of this section is on the former reactions, which are influenced by the full suite of reactions discussed earlier including those that affect (i) solution phase speciation (Section 6.3.2), (ii) adsorption–desorption (Section 6.3.3), and (iii) oxidation–reduction (Section 6.3.4). Mineral weathering in soil is driven by CZ open-system through-fluxes of freshwater and OM, and the resulting introduction of reducing equivalents, complexing ligands, and protons. As a result of mineral transformation reactions, primary minerals (large particles with low specific surface area) are transformed into secondary minerals (small particles with high specific surface area), creating key raw materials for soil fabric, while also releasing lithogenic macro- and micronutrients (e.g., Ca, Mg, K, Fe, Zn, etc.) to solution in support of plant and microbial growth.

#### **6.3.5.2. Soil Biogeochemical Process and Agents**

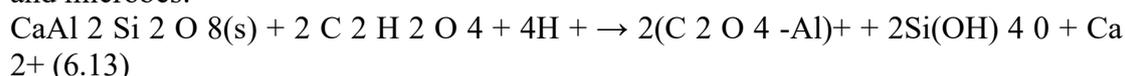
Geochemical weathering reactions in soil are termed congruent if they result in mineral dissolution with no secondary solid formed (i.e., only solution-phase products). Conversely, reactions are termed incongruent if they result in the formation of both solution-phase and solid-phase products. Incongruent weathering reactions are essential to the formation of soils because they result in the production of secondary (pedogenic) minerals from the weathering of primary minerals in rock. Because these secondary minerals are formed under the low temperature (T) and low pressure (P) conditions, most prevalent in the CZ, they tend to have smaller crystallite size and more disordered crystal structure than the primary silicates occurring in the igneous rock undergoing weathering. Their formation here also reflects the greater thermodynamic stability of secondary minerals under low P and T conditions. Secondary minerals do, of course, themselves undergo congruent or incongruent dissolution such that the mineralogical composition of soil evolves over pedogenic time scales (10<sup>2</sup>–10<sup>6</sup> year) in response to climatic forcing and subject to thermodynamic and kinetic constraints.

CZ biota – including plants, microorganisms, and animals – can significantly affect the rates and trajectories of weathering reactions in soils through their impact on the solution-phase reactants – specifically the concentrations of protons, complexing ligands, and oxidizing or reducing equivalents. Atmospherically derived CO<sub>2</sub>(g) is supplemented by root and microbial respiration such that the high partial pressures of CO<sub>2</sub> enhance the thermodynamic favorability of primary silicate weathering reactions. For example, in the following reaction, plagioclase (a framework silicate common to igneous rock) undergoes proton-promoted dissolution in the presence of CO<sub>2</sub> and water to form the layer silicate clay kaolinite:



This reaction reflects several characteristic features of primary mineral weathering in addition to the transformation of primary (in this case framework) to secondary (layer-type) aluminosilicates. These additional characteristic features include water, CO<sub>2</sub>, and/or proton consumption; bicarbonate production; and release of nonhydrolyzing (base) cations, such as Ca, Mg, and K, making them available for bio-uptake from solution or exchange sites. Hence, processes that enhance the chemical activities of water, CO<sub>2</sub>, and protons will increase the thermodynamic driving force – and normally also the kinetics – of rock weathering ( Brantley et al., 2004 ; Stumm, 1997 ; Ugolini and Sletten, 1991 ).

Biologic activity also influences weathering reactions through the production of metal-complexing organic ligands ( Brantley et al., 2011 ; Chorover, 2012 ). These ligands include low molar mass aliphatic and aromatic organic acids, in addition to more specialized high-affinity species (e.g., siderophores). Complexing ligands promote mineral dissolution by sequestering lithogenic metals into metal–ligand complexes (thereby diminishing the solution-phase accumulation of free metals and transporting them from the site of dissolution). For example, the same plagioclase mineral may undergo dissolution in the presence of the common biogenic dicarboxylic acid, ethanedioic or “oxalic” acid (C<sub>2</sub>H<sub>2</sub>O<sub>4</sub>) produced by plant roots and microbes:



By forming a stable complex with Al released from plagioclase, oxalate (the conjugate base of oxalic acid) may preclude the reaction of Al with dissolved Si to form kaolinite, making the dissolution reaction congruent Eq. (6.13) rather than incongruent Eq. (6.12). Furthermore, prior work has shown that low molar mass ligands like oxalate can increase the kinetics of mineral dissolution, even under conditions that are far from equilibrium ( Stumm and Morgan, 1996 ). The mechanism, termed ligand-promoted dissolution, has been shown to be surface-mediated and results from the adsorption of low molar mass ligands like oxalate to surface bound metals via a ligand exchange reaction (similar to Eq. 6.10), followed by dissociation to solution of the metal–ligand complex.

### 6.3.5.3. Reach

The larger the CZ, its geomorphic template, and the local climatic forcing have strong effects on mineral weathering reactions in soils. These CZ-external forcings dictate fluid flow and water–rock contact time, which in turn influence the degree of solution phase saturation with respect to mineral phases undergoing dissolution or precipitation. Such effects, however, can be characterized also in terms of geochemical reactions such as those for plagioclase dissolution or complexing

ligands, reflecting the importance of water flux, proton, and/or CO<sub>2</sub> concentration. These weathering reactions are responsible for generating the secondary (as well as tertiary, quaternary, etc.) minerals in soil environments whose fine particle size and high specific surface area control much of the chemical reactivity of soil systems. They also exert ultimate control over the quality of water emanating from the CZ and entering into surface water bodies. Indeed, the chemistry of surface water is often used in an inverse modeling approach to assess predominating geochemical weathering reactions ( Güler and Thyne, 2004 ; Yoo et al., 2009 ).

6.4. Climate as overarching control on soil geochemistry and its feedback to atmosphere, surface- and groundwater composition Climate is one of the most important forcings impacting CZ evolution ( Anderson et al., 2013 ), soil formation ( Jenny, 1941 ; Volobuyev, 1964 ), and the types/amounts of dissipative products ( Chorover et al., 2011 ; Rasmussen et al., 2011 ). But how is climatic forcing translated to the molecular scale where soil biogeochemical processes are impacted and how do they then feed back to the composition of atmosphere, surface- and groundwater? Because of current challenges of global climate change, impacts of climate on soils is an area of active research ( Kirschbaum, 1995 ; Lal, 2004 ; Peng et al., 2014 ; Schuur et al., 2013 ; Zhang et al., 2013 ) and in many cases temperature and/or precipitation are investigated as climate proxies in isolation. In this section we will give examples illustrating the connection of global/regional climate to molecular scale soil biogeochemical processes that feed back to impact atmosphere, surface- and groundwaters ( Fig. 6.6 ).

#### 6.4.1. Temperature

Temperature has an important bearing on the kinetics of all (bio)geochemical reactions ( Stumm and Morgan, 1996 ). Soil biogeochemical reactions, such as weathering, in “cold” systems may proceed slower than in “warm” systems ( Talibudeen, 1981 ) or may not receive the required activation energy to happen at all. Accordingly, less dissipative products are released to atmosphere and freshwaters in cold systems, than in warmer systems.

Increasing global temperatures have an important effect on the transition from snow to rain ( Elsner et al., 2010 ; Kormos et al., 2013 ) and impact soil biogeochemistry and soil formation. A study by Rasmussen et al. (2007) showed the impact of rain–snow transition on soil development over andesitic bedrock along a climate gradient in the Sierra Nevada of California. Soils in the snowfall-dominated zone (cold) exhibited minimal soil development whereas soils dominated by rainfall (warm) showed the opposite. The authors attribute this pattern to the combined effect of temperature and water availability. In rain-dominated locations, soils are warm when water is available for weathering reactions, whereas in snow-dominated areas soils are cold when water is present.

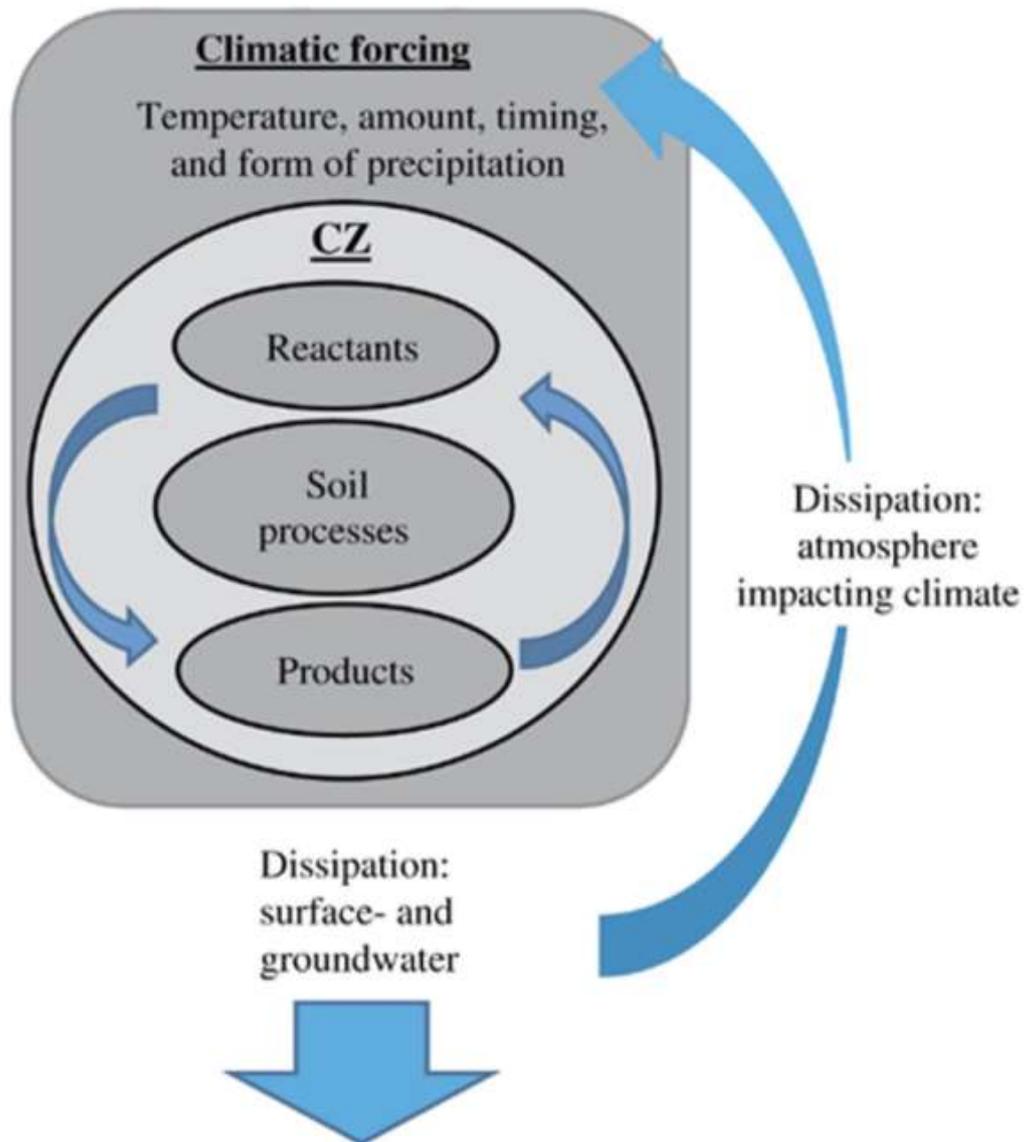


Figure 6.6 Nested factors impacting soil processes and their reach. Climatic forcing acts outside of the Critical Zone (CZ) to impact soil biogeochemical processes that in turn influence composition of atmosphere, surface- and groundwater. Changes in atmospheric composition in turn feeds back to climatic forcings.

#### 6.4.2. Amount and Timing of Rainfall

Soil biogeochemical processes are directly impacted by the amount and timing of precipitation. A certain amount of water is necessary to allow for biogeochemical processes, such as microbial processing, weathering, sorption, exchange reactions, and redox reactions ( Section 6.3 ). For example, soil respiration rates and CO<sub>2</sub> effluxes during the summer show a great sensitivity to moisture, therefore the predicted lengthening of growing seasons and increased summer rains due to global climate change may lead to increased dissipation of C to the atmosphere ( Barron-Gafford et al., 2011 ; Curiel Yuste et al., 2003 ; Fierer and Schimel, 2002 ).

Water delivery also strongly impacts weathering and the dissipation of weathering derived solutes via surface waters ( Gaillardet et al., 2011 ). Loss of solutes at the solid:aqueous interface are controlled by both thermodynamic and

kinetic constraints. For solid phases with slow dissolution kinetics, changes in precipitation may have little influence on the loss of solutes. However, in many catchments, the concentration of solutes in surface- and groundwaters are similar across a wide range of water discharge rates. Catchments falling in this domain are described as chemostatic ( Godsey et al., 2009 ; Maher, 2011 ) and the overall water flux through the CZ controls the mass of solute loss. Maher (2011) demonstrates this well for Si fluxes ( Fig. 6.7 a). The importance of water fluxes for other solute fluxes such as stream water dissolved organic carbon and dissolved inorganic carbon was shown by Perdrial et al. (2014) ( Fig. 6.7 b).

Rain can also induce threshold shifts in soil redox conditions ( Chadwick and Chorover, 2001 ; Miller et al., 2001 ) with potential impacts on solute flux and composition of surface- and groundwaters ( Thompson et al., 2006 ). This is most evident when rainfall exceeds the infiltration capacity of the soil and induces either seasonal or permanent water saturation within the soil profile. Although a greater propensity of anoxic conditions is proportional to rainfall, the response is nonlinear ( Chadwick and Chorover, 2001 ) due to feedbacks between aforementioned secondary mineral production and diminished soil hydraulic conductivity ( Lohse et al., 2009 ; Vitousek and Chadwick, 2013 ). Reductive dissolution of Fe III -solids to Fe 2+ (aq) is evident in the common occurrence of rust colored iron oxide deposits throughout surface waters and springs in high rainfall areas produced as the Fe 2+ (aq) is oxidized when exposed to oxygen ( McKnight et al., 1992 ). Because metal redox reactions produce large shifts in pH (see Section 6.3.4 ), changes in rainfall patterns that alter soil redox conditions can have impacts on streams and ecosystems, potentially contributing to hot spots and hot moments across the landscape ( McClain et al., 2003 ; Small et al., 2012 ).

#### **6.4.3. Amount and Timing of Snowfall and Melt**

Ecosystems that receive an important amount of precipitation in the form of snow (snow dominated) are particularly vulnerable towards global climate change and are increasingly studied (as of 2014, the SoilTrec homepage lists over 60 world Critical Zone Observatories (CZOs) ( <http://www.soiltrec.eu> ), of which at least 1/3 are dominated by a seasonal snowpack). Seasonal snow cover exerts a varied and complex control on soil biogeochemistry because snow dominated systems are (at least) seasonally cold, which has aforementioned implications on geochemical reaction rates. Furthermore, these systems are impacted by the timing of snowfall and water delivery. Soil biogeochemical processes depend on the availability of water in liquid form. In snow-dominated environments, the conversion of water from solid to liquid form represents an important threshold since soil frost is a major inhibitor of soil biogeochemical processes. Snow has a high insulation capacity, therefore soils typically do not freeze completely when protected by a thick snowpack that preserves liquid water for over-winter microbial processes ( Brooks et al., 2011 ; Groffman et al., 2001 ; 2006 ). Such subnivial (under snowpack) processes play an important role in the soils N cycle ( Brooks et al., 1999 ; 2011 ; Monson et al., 2006 ). For example, heterotrophic microbial activity under continuous snow cover reduces N loss allowing for better ecosystem N retention whereas the absence of microbial activity (due to late or missing snow) results in N loss ( Brooks and Williams, 1999 ). Differences in snow cover can also be related to aspect (drier south-vs. wetter north-facing slopes) and can cause differences in solute export from montane environments ( Hinckley et al., 2014 ). Snow cover is equally important for C cycling since more than half of global

organic carbon is stored in high latitude and/or high altitude regions that are at risk to turn from C sinks to sources ( Brooks et al., 2011 ; Tarnocai et al., 2009 ). Estimates of winter CO<sub>2</sub> fluxes due to heterotrophic microbial activity are as high as 60% of the annual net ecosystem exchange ( Brooks et al., 2011 ; Monson et al., 2002 ). Therefore timing and/or amount of winter precipitation is of prime importance for the extent of under snow processes and dissipation of C and N species in gaseous form to impact atmospheric composition or liquid form to influence surface- and groundwaters. It is easy to imagine that in snow-dominated systems, a pronounced disconnect exists between the biogeochemical generation of aqueous dissipative products and their actual dissipation to surface- and groundwaters. Given that liquid water is available at the weathering front, weathering can happen year round but weathering products are only removed during snowmelt. This process leads to the event-like introduction of solutes (weathering products, nutrients, C) ( Boyer et al., 1997 ; Brooks and Williams, 1999 ; Perdril et al., 2014 ) into surface- and groundwaters. Introduction of meltwaters is not only important because of the transporting role of water but also for the reactant role of water. Meltwaters are typically not in equilibrium with mineral phases and therefore chemically aggressive. The introduction of meltwaters under saturated solutions promotes weathering and changes weathering rates in soils over the course of a year. Additionally, meltwaters contain accumulated dust and products of overwinter processes (e.g., subnival microbial processing in organic rich horizons or water–rock interactions at the weathering front) that impact the composition of surface- and low residence time, shallow groundwater.

## **6.5. Conclusions**

As a living porous medium that extends from weathering bedrock to the vegetated surface, soils comprise most of the interfacial area of the CZ. The aggregated structure of bioactive soil contains minerals, microbes, OM, gas, and liquid filled pore spaces all in close association. As a result, soils are characterized by steep geochemical gradients across relatively small time scales (e.g., following hydrologic events) and space scales (e.g., as reflected in distinct redox domains in soil aggregates). Hence, characterizing soil geochemistry tends to be even more challenging than the well-mixed atmospheric and surface water bodies that bound the CZ. The gradients that drive soil geochemical reactions are sustained by continuous open-system throughputs of freshwater and NOM. The ensuing disequilibria result in mineral transformation reactions that favor the formation of high specific-surface-area minerals (layer silicate clays and metal (oxhydr)oxides) with reactive surface sites that contribute to the adsorption–desorption, oxidation–reduction and dissolution–precipitation reactions at highly localized scales. Nonetheless, the importance of these localized reactions is manifested across pedon to catchment to global scales through their regulation of such life-sustaining processes within the CZ, such as carbon stabilization, ecosystem nutrient retention, and contaminant attenuation. As human activities now dominate the structure and function of many portions of the Earth's CZ, the Anthropocene poses tremendous challenges not only to the Earth's surface, but also to researchers engaged in understanding it ( Crutzen, 2006 ; Godd ris and Brantley, 2013 ). Soils and their biogeochemical function are central to mediating changes that occur in this epoch, and it is important to understand how much capacity soil has to absorb changes while still providing life-sustaining services ( Richter and Mogley, 2009 ; Richter and Yaalon, 2012 ).