Rain or Shine

Rain or Shine

An Introduction to Soil Physical Properties and Processes

TYSON OCHSNER

Contents

3. Soil Water Content and Potential

4. Soil Water Flow

5. Water Inputs

6. Infiltration

7. Runoff and Water Erosion

8. Redistribution and Drainage

9. Solute Transport and Groundwater Pollution

10. Evaporation and Wind Erosion

11. Transpiration and Root Water Uptake

12. Surface Energy Balance and Evapotranspiration

13. Soil Temperature

Rain or Shine

An Introduction to Soil Physical Properties and Processes

Tyson E. Ochsner, Oklahoma State University
tyson.ochsner@okstate.edu

Acknowledgements

Cover Image: Rain and shine near Chilcott, Somerset, UK. Photo by Edmund Shaw. Licensed for reuse CC BY-SA 2.0.

Version: 3

Date: January 13, 2019

About the Author

Tyson Ochsner grew up near Chattanooga, Oklahoma, USA, in farming family, whose history is partly described here (link). He earned a B.S. in Environmental Science at Oklahoma State University in 1998. He then studied Soil Science and Water Resources at Iowa State University earning a M.S. in 2000 and a Ph.D. in 2003. From 2003 through 2008, he worked as a soil scientist for the USDA Agricultural Research Service in St. Paul, Minnesota. Since 2008 he has been a faculty member in the department of Plant and Soil Sciences at Oklahoma State University. The mission of his soil physics research and teaching program is to help people better understand and appreciate the soil, the soil water balance, and the surface energy balance so that we can more wisely manage and conserve the land and water with which we have been entrusted. To learn more about his program, please visit http://soilphysics.okstate.edu/. Tyson and his wife, Stephanie, have five children: Audrey, Isaac, Annelise, Eli, and Abel. You can read about one of their great family adventures at this blog (link).

Table of Contents

1. INTRODUCTION

1.1 Purpose and Motivations

"We are, each of us, people of the soil." [1]

The purpose of this book is to serve as a multi-faceted learning resource for people who want or need to learn introductory concepts of soil physics. Soil physics is a scientific discipline focused on using the principles and techniques of physics to understand:

- the properties of the soil,
- the processes which occur in and on soils,
- and how those properties and processes were, are, or would be affected by changes in environment or management.

This book focuses particularly on the processes in the soil water balance and the surface energy balance and how those processes are influenced by soil physical properties.

1.2 Content

In this book, we will begin with consideration of soil patterns, structure, and texture across multiple spatial scales; then walk through the processes of the soil water balance and the surface energy balance, step-by-step. Along the way, we will address most of the soil physics topics included in the Soil Science Fundamentals Exam Performance Objectives defined by the Soil Science Society of America's Council of Soil Science Examiners (available here). Someone who learns the content in this book should be well-prepared for the soil physics portion of that professional examination.

1.3 Audience

This book is primarily written for university students enrolled in an introductory soil physics course at the upper undergraduate levels or at the masters' level. This book is also written for motivated learners anywhere in the world who are curious about the basic concepts and applications of soil physics. This book is designed for students who:

- are comfortable reading and listening in English;
- have basic competence in college-level algebra;
- are familiar with the foundational principles of soil science;
- and understand key physics concepts like mass, energy, force, and pressure.

If you think you may be lacking in one of these areas, don't worry. You can still use this book to learn effectively, but you may need to invest extra effort to fully understand some sections.

1.4 Learning Obstacles and Helps

If you are not comfortable with the foundational principles of math, soil science, or physics, you may benefit from studying resources such as:

- the SSSA Glossary of Soil Science Terms (link)
- Daniels and Haering's "Concepts of Basic Soil Science" (link)
- the "Math Skills Review" by Keeney-Kennicutt (link)
- the Khan Academy video series on Physics (link)

For students in the United States, one of three countries in the world that has not adopted the International System of Units (SI units), this book's use of SI units may be a learning obstacle. The SI units diagram below (Fig. 1-1) provides a helpful visual overview of the seven SI base units and the associated derived units. Derived units that we will use in this book include newtons (N), pascals (Pa), joules (J), watts (W), degree Celsius (°C), and siemens (S).
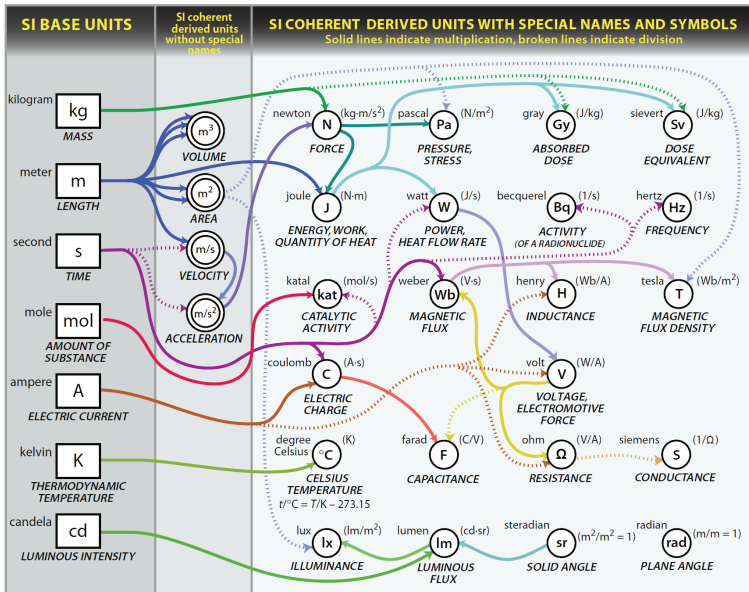
Fig. 1-1. International System (SI) base units and derived units diagram. Created by the US National Institute for Standards and Technology. Reproduced from (link).

To use SI units effectively requires knowing and understanding the SI unit prefixes, which modify the magnitude of the units. The most common SI prefixes are listed in Table 1-1. Being familiar with these prefixes will help you correctly understand the size or intensity of the physical variables we will be studying throughout the book.

Table 1-1. Common SI prefixes. Adapted from Wikipedia (link). CC BY–SA 3.0.

Text	Symbol	Factor	Power
peta	P	1000000000000000	10^{15}
tera	T	1000000000000	10^{12}
giga	G	1000000000	10^{9}
mega	M	1000000	10^{6}
kilo	k	1000	10^{3}
hecto	h	100	10^{2}
deca	da	10	10^{1}
(none)	(none)	1	10^{0}
deci	d	0.1	10^{-1}
centi	c	0.01	10^{-2}
milli	m	0.001	10^{-3}
micro	μ	0.000001	10^{-6}
nano	n	0.000000001	10^{-9}
pico	p	0.000000000001	10^{-12}
femto	f	0.000000000000001	10^{-15}

A related challenge for many students is unit conversion. For example, students in the US often need to convert from customary US units to those of the SI system. Many online unit conversion tools are now available, and in fact, the desired conversion can often be found by typing a phrase like "convert 3.7 inches to centimeters" into the search bar of a web browser. Still it is helpful to know some basic equivalencies by memory, such as 1 inch = 2.54 centimeters (exactly), 1 pound = 0.454 kilograms, and 1 acre = 0.405 hectares (1 hectare is equivalent to 10,000 m2 or 100 m x 100 m). For help with unit conversions, watch this video (link).

There is a common saying in the English language, "It's Greek to me." We use this phrase to mean that we do not understand anything about a given topic. This book does not use Greek language, but it does use Greek letters to represent many of the important constants and variables. Greek letters are often used in

science and mathematics because the 26 letters of the English alphabet may not be enough to uniquely represent all the different constant and variables of interest in a particular area of study and also because of the strong influence of early Greek scholars on the development of Western thought. Learning these Greek letters should make the equations presented throughout this book easier to understand. The Greek alphabet is presented in Table 1-2 for easy reference.

Table 1-2. Greek alphabet.

Uppercase	Lowercase	Name
A	α	alpha
B	β	beta
Γ	γ	gamma
Δ	δ	delta
E	ε	epsilon
Z	ζ	zeta
H	η	eta
Θ	θ	theta
I	ι	iota
K	κ	kappa
Λ	λ	lambda
C	μ	mu
N	ν	nu
Ξ	ξ	xi
O	ο	omicron
Π	π	pi
P	ρ	rho
S	σ	sigma
T	τ	tau
Y	υ	upsilon
Φ	φ	phi
X	χ	chi
Ψ	ψ	psi
Ω	ω	omega

1.5 Problem Set

1. Dimensional analysis

   a. How many millimeters are in 1.31 inch?

   b. The density of a soil sample was 1.47 g/cm^3. This is equivalent to how many kg/m^3?

   c. 1 "acre furrow slice" is a volume of soil approximately 6.0 inches deep occupying exactly 1 acre (43,560 ft^2) of land surface. Estimate the volume in m^3.

2. Manipulation of exponents

   a. $106 \times 102 =$ _____

   b. $(10^2)^{1/3} =$ _____

   c. $(10^{-5} \times 10^3) : 10^2 =$ _____

3. Scientific notations

   a. Write in scientific notations: 0.002413 and 24000

   b. Express $3.937 * 10^{-2}$ as a decimal number

   c. $(7.2 \times 10^5) : (1.2 \times 10^{-2}) =$ _____

4. Logarithms

   a. $Log_{10}\ 50.0 =$ _____

   b. If pH = 7.42, what is the concentration of hydrogen ion?

   c. $\ln (10.) =$ _____

5. Algebraic manipulations. Solve for x.

   a. $5x + 2 = 12$

   b. $(12x + 24)/ 2 = 10\ x$

   c. $(9\ x - 2)/ 3 = (4\ x + 1)/ 2$

6. Significant figures

   a. Give the number of significant figures for: 240, 0.024, 0.0240, 240.0

   b. $2.540 - 1.31 =$ _____

   c. $(2.540 \times 10^2) \times (1.31 \times 10) =$ _____

1.6 References

1. Anonymous. *An Urgent Appeal for Soil Stewardship.* Bouyoucos Conference on Soil Stewardship in an Era of Global Climate Change. 2009. Nebraska City, Nebraska.

2. SOIL PATTERNS, STRUCTURE, AND TEXTURE

Wind and water, gravity and time, people and plants and more change and shape and move the soil, creating patterns and features that range in size from more than 1000 kilometers to less than 2 micrometers. Recognizing and understanding these spatial patterns is vital to more fully understanding and more wisely managing soil physical properties and processes. And, the patterns and structures we can find in soil are fascinating and beautiful. The spatial organization of soil occurs at hierarchical, or nested, levels which we will consider from large to small.

2.1 Global Scale

At the global scale, clear spatial patterns exist in the distribution of soil types as represented by the **soil orders**, the highest level of the USDA soil taxonomy system (Fig. 2-1). The USDA soil taxonomy system (link) is one of several widely-used soil classification systems worldwide (link). The patterns you see in Fig. 2-1 are strongly influenced by the global patterns of climate variables such as temperature and precipitation. In the frigid regions above 60°N latitude, we find large expanses of Gelisols, which are soils with a subsurface layer that remains frozen throughout the year, while the hot and humid equatorial regions of South America and Africa are dominated by Oxisols, the most highly weathered soils on Earth. The east to west gradient of decreasing precipitation across the southern United States is clearly reflected in the progression of soil orders from Ultisols in the southeast to Mollisols in the south central region to Aridisols in the southwest.

These soil order patterns spanning thousands of kilometers also indicate large scale patterns in soil physical properties that influence countless processes in the Earth's coupled human and natural systems. In fact, these global distributions of soil physical properties are not only influenced by climate, but they also influence climate.

Global Soil Regions

Robinson Projection
Scale 1:130,000,000

Soil Orders

Alfisols	Entisols	Inceptisols	Spodosols	Rocky Land
Andisols	Gelisols	Mollisols	Ultisols	Shifting Sand
Aridisols	Histosols	Oxisols	Vertisols	Ice/Glacier
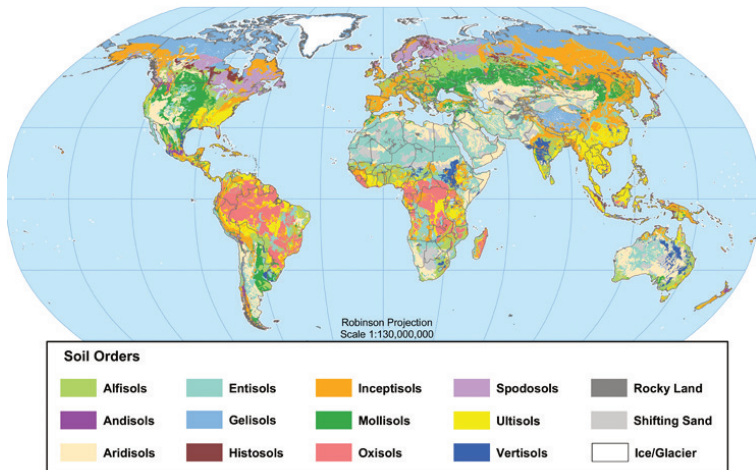
USDA · NRCS · US Department of Agriculture Natural Resources Conservation Service · Soil Survey Division World Soil Resources soils.usda.gov/use/worldsoils · November 2005

Fig. 2-1. Soil orders of the world according to the USDA soil taxonomy system. Source: USDA–NRCS (link).

For example, one significant uncertainty in projections of future climate change is the fate of the large amount of organic carbon stored in the world's permafrost soils [1]. As global temperatures continue to rise, these soils are thawing and the organic carbon is decomposing, resulting in increased carbon dioxide and methane emissions to the atmosphere. If these emissions are not offset by increases in carbon dioxide uptake due to increased plant growth in the warmer climate, then these greenhouse gases released from thawing soils may create a feedback loop that accelerates climate change. Thus, there is a growing need to better understand gas emissions and plant growth in these thawing soils, both of which depend strongly on the soil temperature and moisture dynamics, which are themselves influenced by the soil physical properties.

2.2 Basin or Watershed Scale

At the scale of a river basin or watershed, typically 10s to 100s of kilometers, the spatial patterns in soil properties can be so large that we fail to recognize them. We may only travel a few kilometers from our home in a typical day, or if we travel farther, we may not take the time to notice the changes in the soil around us. Yet, these large spatial patterns can have significant impacts on human communities and on ecosystems.
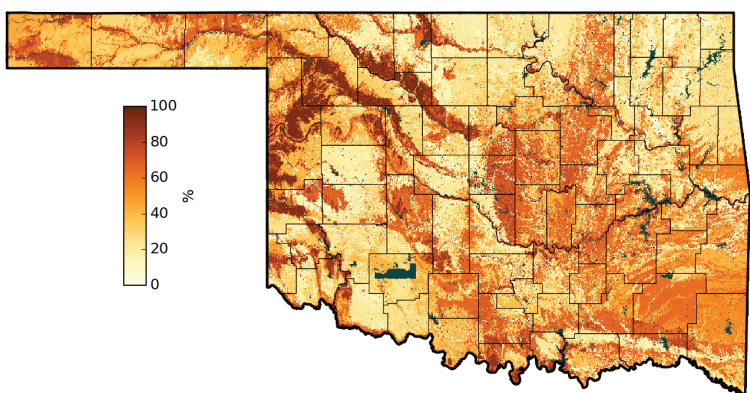
5-cm Sand Content

estimated from USDA-NRCS SSURGO

Fig. 2-2. Large scale movement of wind and water can generate distinctive spatial patterns in soil properties at the river basin scale as shown by this map of sand content of surface soils across Oklahoma, USA. Map credit: Jason Patton.

Consider the map below which shows the estimated sand content of the surface soils across the state of Oklahoma, USA. Some of the striking patterns we can see are the dark brown bands of high sand content trending northwest to southeast in the northwestern part of the state. These sandy areas follow along the north side of rivers such as the Cimarron River and the North Canadian River and

overlie alluvial aquifers which provide groundwater for irrigation and for municipal use by several Oklahoma communities. The windblown sands were likely deposited along the north sides of these rivers by the prevailing south winds during past severe drought events which denuded the landscape. These soil patterns help explain a variety of social and ecological characteristics from the economies of rural communities to the distributions of native plant species on the landscape.

2.3 Hillslope Scale

One level below the watershed scale, we find soil spatial patterns at the hillslope scale, roughly 10s to 100s of meters in size. Erosion and deposition help create many of the spatial patterns at this scale. Classic examples are alluvial fans, which are fan-shaped deposits of water transported sediment at the base of hills or mountains. The photograph below provides an aerial view of a well-developed alluvial fan in the French Pyrenees mountains. If you look carefully, you can see the trails perpendicular to the slope created by grazing livestock, indicating focused zones of soil compaction.

Fig. 2-3. Erosion and deposition processes contribute to spatial organization of soils at the hillslope scale as shown by this alluvial fan in the Cirque d'Estaubé, French Pyrenees. 42° 42′ 50.4″ N, 0° 02′ 56.4″ E. Photo credit: Mikenorton CC-BY-SA 3.0.

Another important hillslope scale pattern commonly recognized by soil scientists is called a catena or toposequence. A catena is a series, or chain, of distinct but related soil types arranged along a hillslope. The soils in a catena typically form from the same parent material but have different physical properties arising from differences in slope, aspect, and drainage. The diagram below represents the Clarion–Nicollet–Webster catena, a common catena found in the highly productive farmland of Iowa and Minnesota in the midwestern United States.
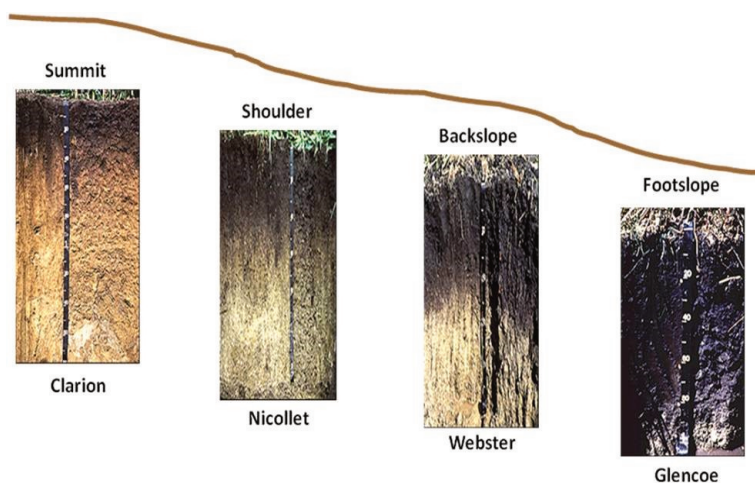
Fig. 2-4. The solid line represents the land surface topography and the photos show the respective locations of the soil series in a Clarion-Nicollet-Webster catena. The soil color varies from red/brown subsurface colors at the summit, indicating oxidizing conditions, to gray subsurface colors at the footslope, indicating reducing conditions. Image credit: Jay Bell, Univ. of Minnesota. CC-BY-NC-SA.

These soils formed in unsorted parent material deposited by prior glaciers, called till, and they differ in drainage with the Clarion soils on the summits being moderately well drained, the Nicollet soils on the shoulder slopes being somewhat poorly drained, and the

Webster soils on the backslopes being poorly drained. The Glencoe soils are more limited in extent, very poorly drained, and found in closed depressions. These differences in drainage impact a range of soil properties, for example, soil color and thickness of the A horizon, as seen in the profile photographs.

As a result of these hillslope scale soil patterns, many farmers in this region have installed elaborate subsurface drainage systems which remove water from the poorly drained portions of the field in order to improve crop growth. The complex drainage system mapped below drains approximately 129 hectares in western Minnesota, i.e. 320 acres or two quarter sections in the US public land survey system [2]. Notice that many of the drainage pipes have been placed in the low areas and depressions, locations occupied by the more poorly drained soils in the local catena. In this case those soil series are the Parnell and Flom series, both classified as aquolls according to the US soil taxonomy system. The artificial drainage of soils like these greatly increases the crop production potential of the land, but also decreases wetland habitat and increases the risk of contaminating the water bodies receiving the drainage water. To learn about improved drainage water management practices, view this video (link).
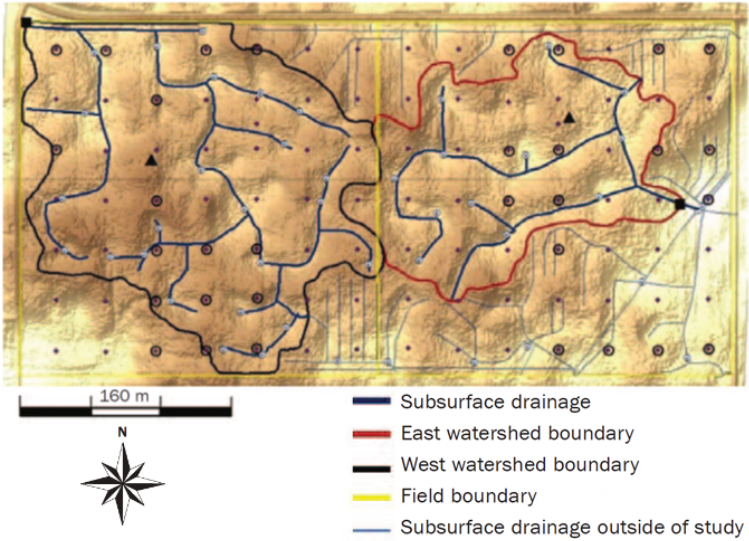
Fig. 2-5. A large subsurface drainage system under cropland in western Minnesota, USA. The thick blue lines and thin blue lines show the locations of underground drainage pipes which vary in diameter from 10 to 61 centimeters (4 to 24 inches). The topography is indicated by background shading, with lower areas being lighter colored. The dots, circles, and triangles mark features relevant to a research study conducted at this site. Adapted from Krueger et al. (2013).

2.4 Soil Profile Scale

Below the hillslope scale, soil profiles display vertical spatial organization and patterns spanning 10s to 100s of cm. The most significant patterns at this scale are the layers, or horizons, which together make up the soil profile and differentiate one soil type from another. Adjacent horizons differ from each other in physical and/or chemical properties, and the specific combination and sequence of horizons in a soil profile influence, and are influenced by, the presence and passage of water, energy, and living organisms. A soil profile of the Tetonka soil series shown in Fig. 2-6 provides a dramatic example of the stark contrasts which can occur between adjacent horizons. The first, or uppermost, horizon is a >30-cm thick A horizon that is dark gray to black due to its high soil organic matter content. Just below that is an E horizon with a strongly contrasting light gray to white color resulting from the downward leaching, or eluviation, of clays, iron oxides, and aluminum oxides. Below the E horizon, is the Bt horizon which is marked by accumulation of clays leached from above. If you look closely, you can see evidence of prismatic type soil structure in this Bt horizon, suggesting the presence of yet another underlying level of spatial organization.

Fig. 2-6. Striking soil horizon differentiation in a grassland soil in South Dakota, USA (Tetonka series). The scale is marked in feet. Photo Credit: University of Idaho (link).

2.5 Soil Aggregate Scale

When we look closer still, we can see that the soil solids are often organized into three-dimensional assemblages ranging in size from approximately 0.1 to 10 cm. These assemblages are called aggregates, and their presence and characteristics are the defining features of what is commonly called **soil structure**. Small, rounded, highly-porous aggregates are indicative of soil structure that is well-suited for plant growth; whereas, large, angular aggregates with low porosity indicate structural conditions that are likely to limit plant growth. Often,a wide array of different aggregate sizes are present in a soil at the same time, as illustrated by the photo below. Soil structure and soil aggregates can change over time and are often altered, for better or worse, by human management practices such as tillage, crop planting, or preparing soil for construction projects. Properly-timed tillage practices using suitable implements can improve soil structural conditions near the surface, but these improvements are often temporary. Poorly-timed tillage or traffic by equipment or livestock can degrade soil structure resulting in surface or subsurface compacted layers that may last indefinitely.

Fig. 2-7. Soil aggregates with a wide range of sizes constitute the surface layer of this clay-rich soil found near Auroville, Tamil Nadu, India. Photo credit: Carla Antonini, CC–BY–SA–3.0 (link).

Soil aggregate size distributions can be measured while the aggregates are either wet or dry, depending on the purpose of the measurement. Researchers often use a stack of sieves to measure the size distribution of aggregates remaining intact after soaking and shaking initially dry aggregates of a specific size range (e.g. 4-8 mm) in a dilute solution. This process gives the size distribution of **water stable aggregates** which is intended to represent a soil characteristic called **aggregate stability**.
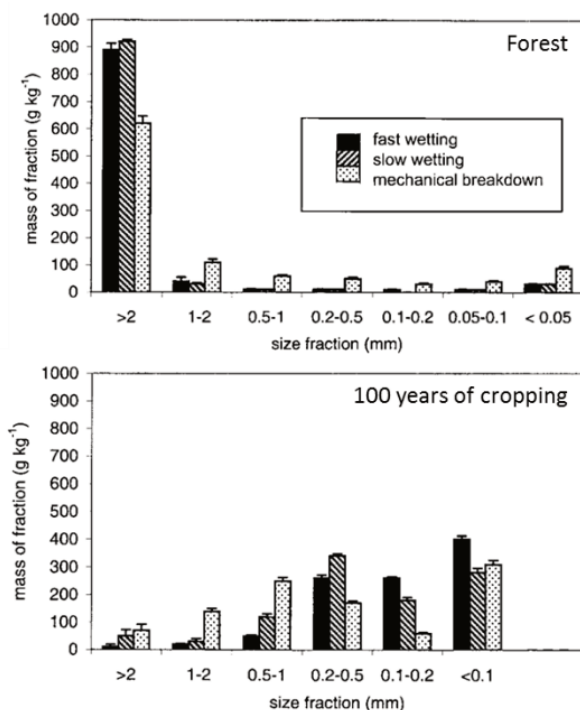
Fig. 2-8. Aggregate size distributions after water stability tests for two soils in France. Reproduced from Chenu et al. (2000).

Fig. 2-8 shows size distributions of water stable aggregates for two soils in France [3]. The soil from the forest had greater aggregate stability as shown by the fact that the >2 mm size fraction dominates the distribution. The cropland soil had lower aggregate stability as shown by the larger mass fractions in the smaller size classes (<2 mm), and data from other sites showed that the aggregate stability was lower the longer the soil had been farmed. This trend of decreasing aggregate stability was likely due to a similar decreasing trend of soil organic carbon which varied from 52.6 g kg^{-1} for the forest soil to 8.4 g kg^{-1} for the soil which had been farmed for 100 years. These data illustrate the general tendency for aggregate

stability to increase as soil organic carbon increases. Soils with higher aggregate stability are typically less susceptible to surface crusting and erosion by natural rainfall, topics addressed in a later chapter.

Fig. 2 9. Simple method for demonstrating aggregate stability. Photo credit: USDA-NRCS (link).

For educational purposes, a simpler procedure can be used to demonstrate aggregate breakdown due to wetting or to show differences in aggregate stability between soils (Fig. 2-9). For this photograph, two initially dry aggregates from different soils were immersed in water. The aggregate on the left is more stable in water and thus loses less soil than the aggregate on the right. The darker color and greater stability of the aggregate on the left suggest that it has higher soil organic matter content than the one on the right.

Aggregate size distributions are also commonly measured by sieving air-dried soil samples, and the results are intended to

represent the aggregate sizes existing in the field at the time of sampling, apart from the influence of rain or water. One way of summarizing the aggregate size distribution of a soil is by calculating the geometric mean diameter (GMD) using

$$GMD = \exp\left[\left(\sum_{i=1}^{n} W_i \ln \bar{X}_i\right) \Big/ \sum_{i=1}^{n} W_i\right]$$

(Eq. 2-1)

where W_i is the mass of aggregates in size class i of n total size classes each with average diameter . The GMD provides a single number which can be used to compare the aggregate size distributions between various soils.

The dry aggregate size distribution of the surface soil can influence the suitability of the soil as a seedbed for crops and the susceptibility of the soil to wind erosion. For example, the time from planting to emergence of corn seedlings was influenced by the aggregate size distribution with earliest emergence occurring when the GMD was between 1.0 and 6.8 mm during growth chamber experiments using soils from the Midwestern US (Fig. 2-10, [4]). The best tillage and planting procedures should be the ones that result in GMD values in the optimal range, which may be soil specific.
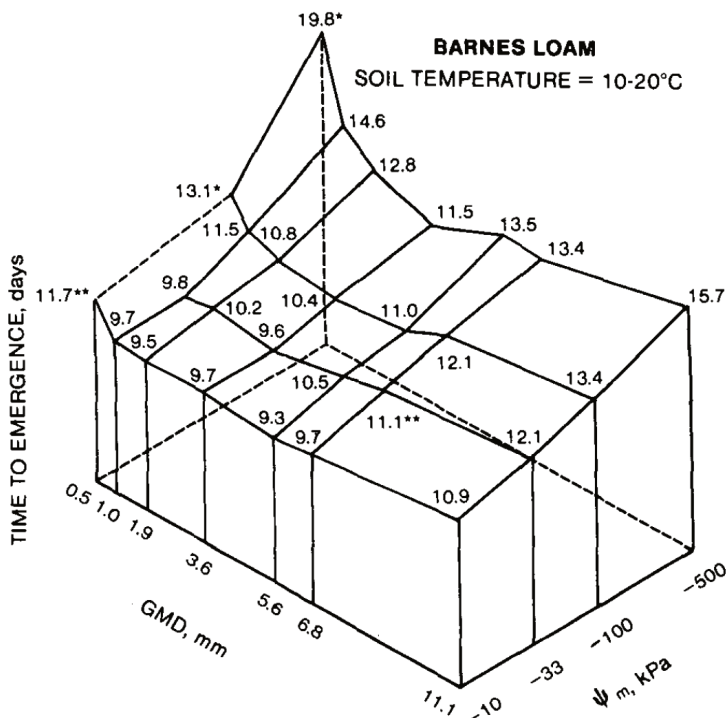
Fig. 2-10. Days from planting to emergence for corn in a loam soil as influenced by geometric mean diameter (GMD) of the soil aggregates and soil matric potential (ym). Soil temperatures varied from 10–20°C daily. Adapted from Schneider and Gupta (1985).

However, most farmers do not have easy access to the equipment needed to evaluate the aggregate size distribution of their soil, so GMD may be of limited practical value in evaluating soil management practices. More practical approaches for land managers utilize simple tools such as spades and rulers and rely on human sight and touch. One such method which has proven useful in several nations around the world (Fig. 2-11) is called Visual Evaluation of Soil Structure [5]. In this method, the evaluator removes an intact slab of soil spanning roughly the top 20-cm,

follows a step-by-step process to carefully observe the structure of that slab in the field, and then assigns a soil structural quality score to either the whole slab or noticeable layers within the slab using a reference card with photographic examples and key diagnostic criteria. (If you print this card for your own use, note that the measuring scale along the edge may not be accurate due to automatic rescaling of the page to your paper size.) In these instructive videos, Dr. Bruce Ball and Dr. Rachel Guimarães demonstrate the visual evaluation of soil structure for a compact soil and a loose soil. The soil structural quality scores resulting from this procedure have proven to be more strongly correlated with crop yields than are more traditional quantitative measures of soil physical properties such as penetration resistance and bulk density [6].

Fig. 2-11. Tyson Ochsner demonstrating visual evaluation of soil structure to farmers in Haiti. Photo credit: Shane Robinson.

2.6 Scale of Primary Soil Particles

When we break down the soil aggregates completely, we are left with the individual soil particles, and the size distribution of these primary particles defines the **soil texture**. Typically, we classify these particles as either sand, silt, or clay based on their size. In the widely-used classification system of the United States Department of Agriculture, particles < 2.0 and ≥ 0.05 mm are considered **sand**, particles between < 0.05 mm and ≥ 0.002 mm are considered **silt**, and particles < 0.002 mm are considered **clay**. The **particle size distribution** or soil texture is one of the most significant soil physical properties, and it has a major influence on the chemical, physical, and biological processes that occur in the soil.

To measure the particle size distribution, we typically disperse, or separate, the soil particles in a liquid to create a suspension using either chemical or physical dispersion methods, or both. Sometimes pre-treatment of the sample is necessary to remove binding agents such as soil organic matter. Once the soil particles are completely dispersed and well-mixed in the suspension, the particle size distribution can be determined based on the rate at which particles settle out of the suspension using sedimentation theory. Alternatively, the particle size distribution can be determined using laser diffraction methods, which can also provide information about the particle shape [7].

Sedimentation, or sediment deposition, is the process of particles settling out of suspension, and it is the process most commonly used to determine soil particle sizes and soil texture. You have probably observed that solid particles in a fluid settle at different rates depending on their size and density. In soil science, we often use **Stokes' Law** to represent the relationship between a particle's size, its density, and its settling velocity. For a spherical particle a

conceptual diagram of the situation looks like Fig. 2-12. The symbol, F_g, represents the downward force due to the difference between the weight of the sphere and the buoyancy of the sphere:

$$F_g = \frac{4}{3}\pi r^3 (\rho_s - \rho_f)g$$

where r is the radius of the particle, ρ_s is the particle density, ρ_f is the fluid density, and g is the acceleration due to gravity.
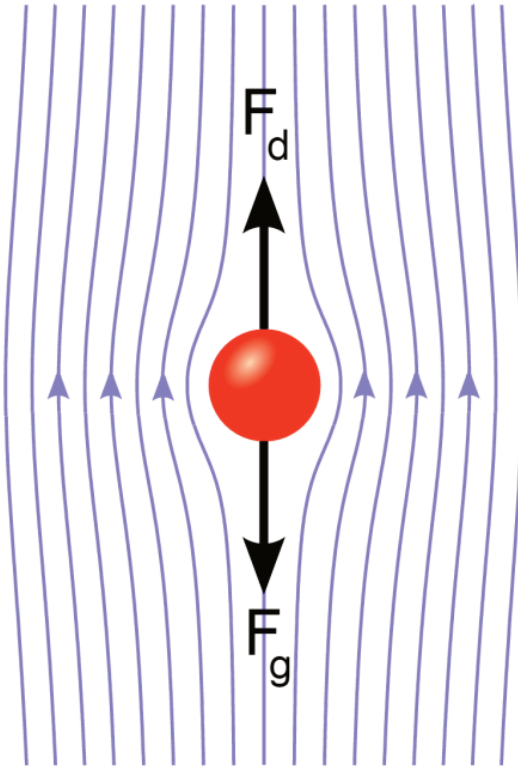
Fig. 2-12. Conceptual diagram of a spherical particle settling out of a fluid. Image credit: Kraaiennest, CC BY–SA 3.0 (link).

When the falling particle is no longer accelerating, it has reached terminal (or maximum) velocity, and the downward force due to gravity is perfectly balanced by an upward drag force, F_d, due to the particle's motion through the fluid:

$$F_d = 6\pi\eta r u$$

where h is the dynamic viscosity of the fluid and u is the terminal velocity. The **dynamic viscosity** is basically a fluid's resistance to flow and is similar to what might commonly be referred to as the "thickness" of the fluid. Setting $F_g = F_d$ and rearranging, we can solve for the terminal velocity:

$$u = \frac{d^2 g(\rho_s - \rho_f)}{18\eta}$$

(Eq. 2-2)

where d is the particle diameter. Once we know the settling velocity for particles of a given diameter, we can easily calculate the time required for those particles to fall a specified distance through the fluid because velocity is simply distance divided by time. That settling time calculation guides the sampling time used in laboratory methods for determining particle size distribution such as the hydrometer method and the pipette method [8]. For a video explaining and applying **Stokes' Law**, click here.

When we use Stokes' Law to determine the soil particle size distribution, we typically assume that:

- the soil particles are shaped like spheres
- the suspension is dilute enough that the particles do not interact with each other
- the motion of the falling particles is slow enough that it does not generate turbulence
- the particles reach terminal velocity immediately after we stop stirring the suspension

The accuracy of our results then depends on the validity of these assumptions. For example, although sand particles might be roughly

approximated as spheres, clay particles resemble something more like thin plates (Fig. 2-13). Although the shape asumption and others may be violated in standard soil particle size analysis methods, still the application of Stokes' Law for this purpose has proven to produce reliable and consistent particle size information that is foundational for our physical understanding of soils. If you are curious, you can see the settling behaviour of particles that drastically violate Stokes' Law demonstrated in this fascinating high speed video (link).
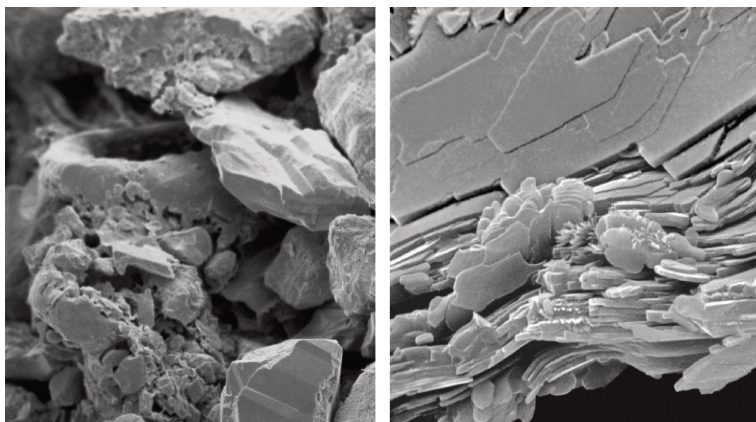
Fig. 2 13. Scanning electron microscope pictures of soil sample consisting of primarily sand-sized particles (left) and a sample of illite clay (right). The clay is shown at higher magnification than the sand. Image credit: Eickhorst, Thilo & Tippkoetter, Rolf. Micropedology – The hidden world of soils. University of Bremen, Germany (link). CC BY-SA 3.0.

Once we have determined the percentages of sand, silt, and clay in a soil sample, we can identify the soil textural class appropriate for that sample by following the USDA classification scheme or other appropriate national or international schemes. Knowing the sand, silt, and clay percentages, we can use a soil textural triangle to classify a sample (Fig. 2-14). In textural class names, the last word is the primary classifier and the preceding word or words provide

additional descriptive detail. For example, a silty clay is clayey soil with a relatively large amount of silt. In addition to sand, silt, and clay, the words loam and loamy are used to indicate a mix of all three particle size classes.
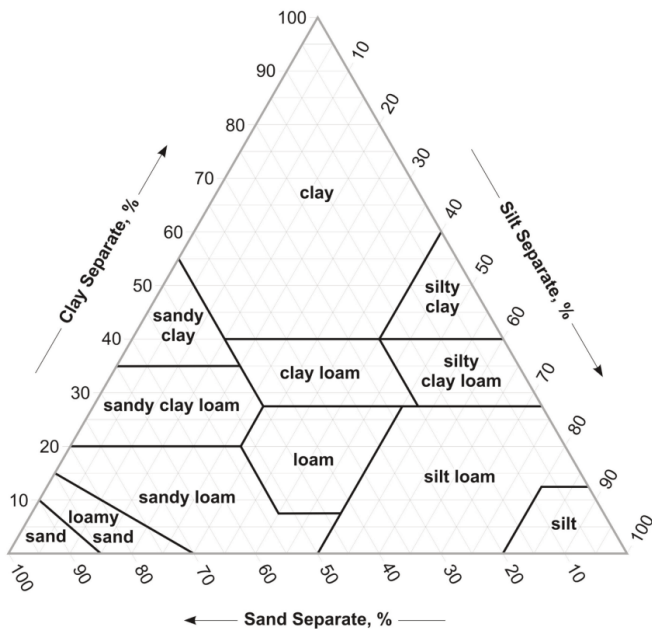
Fig. 2-14. Soil textural triangle showing the 12 USDA soil textural classes. Read the percent sand along the lines sloping up and to the left from the bottom axis, the percent clay along the horizontal lines, and the percent silt along the lines sloping down and to the left from the axis on the right hand side. Source: USDA–NRCS (link).

Sometimes we want to estimate soil texture in the field or without the time and expense of the laboratory procedures discussed above. In that case, one can learn, with practice, to estimate the soil textural class accurately by touch and feel, using only a small amount of water. The silent video here and the flow chart below

(Fig. 2-15) describe a widely-used procedure for estimating soil texture by feel according to the USDA soil textural classes [9].
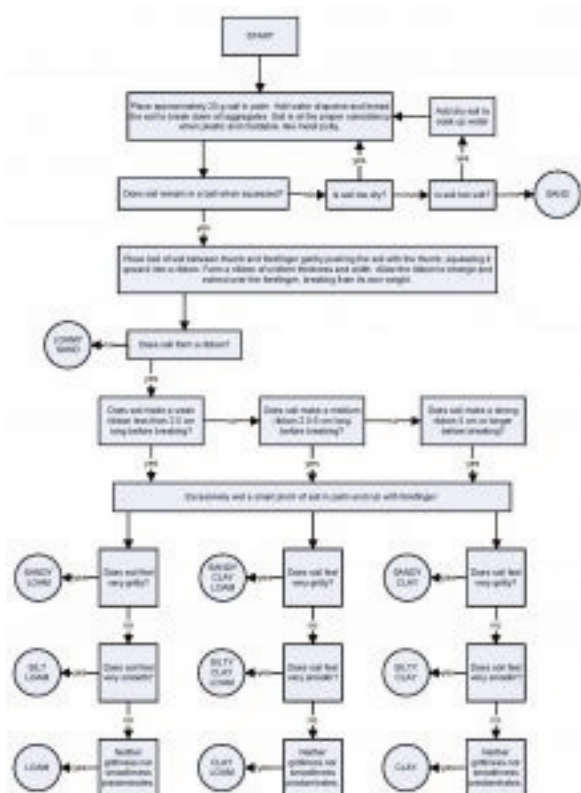
Fig. 2-15. Flow chart for estimating USDA soil textural class by touch and feel. Source: USDA–NRCS (link).

Imagine if you wanted to understand or predict the transport of a pesticide through the soil in Fig. 2-16 and the pesticide was prone to sorption by the clay. The actual transport behavior might be dominated by water flow through the large pores and by sorption of the pesticide to the clay coatings along those pores, and might be dramatically different from the behavior you would observe if

you homogenized the sample and destroyed the micro-scale spatial organization. Researchers have often used disturbed and homogenized soil samples to study soil properties and processes in the laboratory, but in part because of advances in high resolution soil visualization technologies, the strong influence of micro-scale soil organization and the necessity to study the soil intact are becoming increasingly clear.Even when considering soil at the scale of the primary soil particles, we need to recognize that soil is not a homogeneous mixture of particles. Even at this scale where distances are measured in microns, intricate and influential spatial patterns and spatial organization are the rule, not the exception. Consider the beautiful microscopic image below of clay coatings along the pore walls inside a soil from Italy (Fig. 2-16). Coatings like these typically result from the transport of clay particles by water flowing through the soil pores and deposition of those particles along the pore walls.
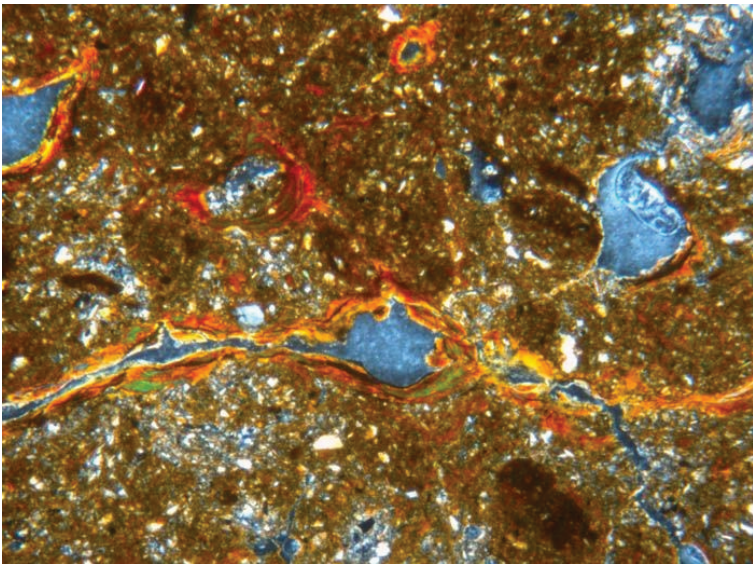
Fig. 2 16. Microphotography of a clay coating in an Italian Alfisol.
Cross-polarized light; frame width is 1 cm. By Francesco Malucelli, CC
BY-SA 2.5 (link).

2.7 Terminology

As you continue to study soil physical properties and processes, you will need to know and use the following key terms related to soil structure and texture:

- **Bulk density** (ρ_b) is the mass of oven-dry soil solids per unit volume of soil. In this context, soil is oven-dry when it has been dried at 105°C to a constant weight. Bulk density has dimensions of mass over volume and units of $\mathrm{Mg\ m^{-3}}$, $\mathrm{kg\ m^{-3}}$, or $\mathrm{g\ cm^{-3}}$. For mineral soils it ranges from approximately 0.80 to 2.0 $\mathrm{g\ cm^{-3}}$. Organic soils may have lower values. Bulk density is distinct from wet bulk density which is sometimes used in engineering contexts.
- **Particle density** (ρ_s) is the mass of oven-dry soil particles per unit volume of soil particles. The volume of soil particles is smaller than the bulk soil volume, which includes pore spaces. Therefore, particle density is greater than bulk density. Particle density has dimensions of mass over volume and units of $\mathrm{Mg\ m^{-3}}$, $\mathrm{kg\ m^{-3}}$, or $\mathrm{g\ cm^{-3}}$. For mineral soils a value of 2.65 $\mathrm{g\ cm^{-3}}$ is often assumed if the true particle density is unknown.
- **Porosity** (ε) is the volume of pore spaces per unit volume of soil. Porosity has dimensions of volume over volume and may be written as a unitless decimal, as a percentage, or as a decimal with units of $\mathrm{m^3\ m^{-3}}$ or $\mathrm{cm^3\ cm^{-3}}$. We will use the latter in this book. If bulk density is known, then porosity can be estimated using the following equation. To better understand the equation, watch this brief video.

$$\varepsilon = 1 - \frac{\rho_b}{\rho_s}$$

(Eq. 2-3)

2.8 Problem Set

1. Calculate the terminal velocity in water at 20°C for each of the following. Assume a particle density of 2650 kg m^{-3}, a water density of 1.00×10^3 kg m^{-3}, and a viscosity of 1.00×10^{-3} kg m^{-1} s^{-1}.

   1. 100. μm diameter sand particle
   2. 10.0 μm diameter silt particle
   3. 1.00 μm diameter clay particle

2. Determine the USDA textural class for the following soil samples:

Sample #	Sand (%)	Silt (%)	Clay (%)	Textural class
1	26	42	32	
2	85	7	8	
3	33	54	13	

3. Go outside and use texture by feel to find soil from two different textural classes and report the latitude, longitude, and estimated textural class. If you are in the USA, use the SoilWeb app (iOS or Android) to determine and report the expected soil series for your location. Do not trespass. Use all appropriate safety precautions.

4. The following 5.0 cm diameter soil samples that are 5.1 cm in length were dried at 105°C for 24 hours. Calculate the bulk density and porosity of each sample.

Sample #	Dry Mass (g)	Bulk Density (g cm^{-3})	Porosity (cm^3 cm^{-3})
1	168		
2	105		
3	132		

2.9 References

1. Koven, C.D., et al., *A simplified, data-constrained approach to estimate the permafrost carbon–climate feedback.* Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2015. **373**(2054).

2. Krueger, E., et al., *On-farm environmental assessment of corn silage production systems receiving liquid dairy manure.* Journal of Soil and Water Conservation, 2013. **68**(6): p. 438-449.

3. Chenu, C., Y. Le Bissonnais, and D. Arrouays, *Organic matter influence on clay wettability and soil aggregate stability.* Soil Science Society of America Journal, 2000. **64**(4): p. 1479-1486.

4. Schneider, E. and S. Gupta, *Corn emergence as influenced by soil temperature, matric potential, and aggregate size distribution.* Soil Science Society of America Journal, 1985. **49**(2): p. 415-422.

5. Guimarães, R.M.L., B.C. Ball, and C.A. Tormena, *Improvements in the visual evaluation of soil structure.* Soil Use and Management, 2011. **27**(3): p. 395-403.

6. Munkholm, L.J., R.J. Heck, and B. Deen, *Long-term rotation and tillage effects on soil structure and crop yield.* Soil and Tillage Research, 2013. **127**: p. 85-91.

7. Zobeck, T.M., *Rapid soil particle size analyses using laser diffraction.* Applied Engineering in Agriculture, 2004. **20**(5): p. 633-640.

8. Gee, G.W. and D. Or, *2.4 Particle-size analysis*, in *Methods of Soil Analysis: Part 4 Physical Methods*, J.H. Dane and C.G. Topp, Editors. 2002, Soil Science Society of America: Madison, Wisconsin. p. 255-293.

9. Thien, S.J., *A flow diagram for teaching texture-by-feel analysis.* Journal of Agronomic Education, 1979. **8**(2).

3. SOIL WATER CONTENT AND POTENTIAL

In the prior chapter we considered the multi-scale patterns, structure, and texture associated with the soil solid phase. In this chapter we will turn our attention to the soil liquid phase, which we will simply refer to as soil water in most cases. It is important to keep in mind however, that the soil liquid phase is not simply water, but rather a complex solution containing organic and inorganic solutes, not to mention the micro-organisms and colloids often suspended in the solution. The soil water is in some ways analogous to a river transporting life-giving, and sometimes life-threatening, solutes, organisms, and colloids. But those are topics for another chapter. Here we will focus on the amount and the condition of soil water.

3.1 Soil Water Content

When the United States National Aeronautics and Space Administration (NASA) was building the Phoenix Mars Lander to explore the surface of Mars, the project leaders contacted soil physicists to create a special sensor for the rover's robotic arm (Fig. 3-1). The purpose of that sensor was to test for the presence of water in the Martian soil. The fascinating story of the resulting partnership between NASA and soil scientists is recounted in this video. The NASA scientists knew that finding water on Mars was key to answering questions about the possibility of life on that hostile planet. Water is the life-blood of the soil. Almost all terrestrial life and almost all soil biological, chemical, and physical processes are influenced by the water content of the soil.
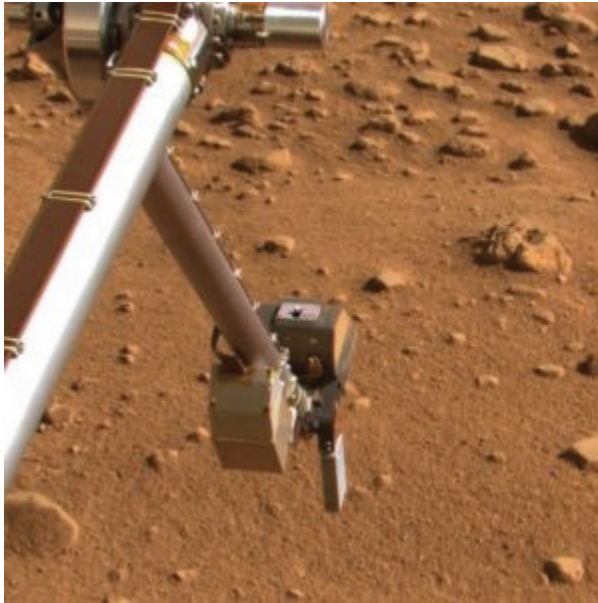
Fig. 3-1. NASA's Phoenix Mars Lander's Robotic Arm with its thermal and electrical conductivity probe, at the lower right, pointing toward the Martian soil (July 14, 2008). Photo credit: NASA/JPL (link).

3.1.1   Soil Water Content Terminology

As you continue to study soil physical properties and processes, you will need to know and use the following key terms related to soil water content.

- *Gravimetric water content* (θ_g) is the mass of water per unit mass of oven-dry soil particles. It has dimensions of mass over mass and may be written as a unitless decimal, as a percentage, or as a decimal with units of kg kg^{-1} or g g^{-1}. We will use the latter in this book. For mineral soils it ranges from approximately 0 to 0.60 g g^{-1}. Organic soils may have higher

values.

- **Volumetric water content** (θ) is the volume of water per unit volume of soil. It has dimensions of volume over volume and may be written as a unitless decimal, as a percentage, or as a decimal with units of $m^3\ m^{-3}$ or $cm^3\ cm^{-3}$. We will use the latter in this book. For mineral soils it ranges from approximately 0 to 0.60 $cm^3\ cm^{-3}$. Organic soils may have higher values. When gravimetric water content is known, the volumetric water content can be calculated using the soil bulk density (ρ_b) and the density of water (ρ_w). This relationship is further explained in this video.

$$\Theta = \Theta_g \frac{\rho_b}{\rho_w}$$

(Eq. 3-1)

Soil water storage (S) is the equivalent depth of soil water contained in a soil layer of a specified thickness (Δz). It has dimensions of length and may be written as cm or mm. If the volumetric water content of the layer is known, then the soil water storage is calculated by S = θ × Δz.

3.1.2 Soil Water Content Measurement Techniques

Knowing the amount of water contained in the soil is important, not only in the search for life on Mars, but also for understanding and managing life on Earth. Researchers in disciplines such as soil science, hydrology, ecology, agronomy, meteorology, and civil engineering often need accurate measurements of soil water content; as do farmers, horticulturalists, turf managers, and government agencies responsible for managing natural resources. Since around the end of the World War II, researchers have been

continually developing new methods for measuring soil water content to meet the needs of these diverse user groups [1]. We will briefly consider a few of those measurement techniques here.

The simplest and oldest method for measuring soil water content is called the **gravimetric method**. In this method, a soil sample is weighed, then oven-dried at 105°C to a constant weight and the final dry weight is recorded. This drying temperature is above the boiling temperature for water but below temperatures at which soil organic matter would be lost by heating. If the volume of the soil sample is unknown, then only the gravimetric water content can be determined using this method. If the volume of the sample is known, then the bulk density and volumetric water content can also be determined. In practice, a fixed drying time of 24 hours is often used because it is more convenient than repeatedly weighing the sample to determine if a constant weight has been reached. The primary requirements for this method then are an accurate balance and an oven that can safely hold a constant temperature unattended for 24 hours.

In contrast to the simplicity of the gravimetric method, powerful new methods of measuring and monitoring soil water content involve a diverse array of cutting-edge technologies. Now satellites orbit the Earth carefully mapping the emission of electromagnetic radiation from the land surface in the microwave wavelength range, and these data are used to estimate the volumetric water content of the 0-5 cm soil layer around the globe with ~9 km spatial resolution [2, 3]. Researchers are also learning to extract information about soil water content from the satellite-generated signals received by antennas of Global Positioning Systems because a portion of those signals comes from reflections off the land surface [4, 5].

Meanwhile, neutrons generated by cosmic-rays are detected with stationary or mobile neutron counters at the land surface and used to estimate volumetric water content of the surface soil to a depth of up to 40 cm with a spatial resolution of ~400 m [6, 7]. For higher spatial resolution, researchers deploy fiber optic distributed temperature sensing systems with or without active heating

capabilities. These systems can measure soil temperatures at 1-m resolution and 1-min intervals along paths of 1-km length or greater, and the soil temperature data can be used to infer soil volumetric water content [8, 9]. Fig. 3-2 shows some of the equipment in a field experiment called the Marena, Oklahoma, In Situ Sensor Testbed (MOISST) where many of the established and emerging soil moisture measurement technologies are inter-compared [10].

Fig. 3-2. Sensors at the Marena, Oklahoma, In Situ Sensor Testbed (MOISST) site including sensors for the cosmic-ray neutron method (white box on right) and GPS reflectometry method (gray domes on masts), as well as numerous types of below ground soil water content sensors.

Add to these techniques the steady progression of new electromagnetic and thermal sensors that permit long-term, automated monitoring of soil volumetric water content at a single point [11] , a progression that began with the seminal work of Topp et al. [12] who used a technique called time domain reflectometry (TDR). You can see that development of improved methods for measuring soil water content has been a major focus in soil physics and related disciplines, and you can understand why soil water

content data are becoming increasingly available around the world [13]. One of the great opportunities for researchers today is to find creative ways to use these data for societal benefit whether for improved drought monitoring, more accurate streamflow forecasting, increased wildfire preparedness, or a host of other potential applications [14-16].

3.2 Soil Water Potential

Knowing the soil water content is useful for many applications, but there is another variable which is equally important to understanding soil water processes, and that variable is called soil water potential. Soil water potential is a measure of the potential energy of the soil solution. **Potential energy** is energy that a substance or object has by virtue of its location or internal condition, and potential energy is always defined relative to some reference condition. For example, a book resting on a table has less potential energy than the same book would have if you held it in the air above the table. If you release the book, it will spontaneously fall back down to the table due to the force of gravity. In this case, the increased potential energy came from raising the book upward against the Earth's gravitational field, and the book's spontaneous fall when released illustrates the general tendency of things to move from locations of higher to lower potential energy.

3.2.1   Soil Water Potential Terminology

Thus, **soil water potential** (ψ) is the potential energy of the soil solution relative to that of water in standard reference state. By convention, that state is defined by:

- pure water
- at atmospheric pressure
- at the same temperature as the soil solution
- at a specified, constant elevation.

Differences in soil water potential cause water in soil to flow from regions of higher water potential to regions of lower water potential. When the water potential is equal throughout a certain

region of soil, then that region has reached a condition called **equilibrium**. At equilibrium, there is no water flow. The concept of equilibrium is important for theoretical purposes and laboratory experiments but rarely, if ever, occurs in the field.

Most scientists consider the total soil water potential to be the simple sum of component potentials from several contributing factors, most notably: gravitational potential, osmotic potential, and pressure potential.

- **Gravitational potential** (ψ_g) is the potential energy which is due to the position of substance or object in a gravitational field. Gravitational potential is measured with respect to an arbitrary reference elevation.
- **Osmotic potential** (ψ_o) is the decrease in the water potential which is due to the presence of solutes. Osmotic potential affects soil water flow whenever there is a gradient in solute concentration within the soil. Such gradients are especially persistent where there is a membrane, such as the wall of a plant cell, or a barrier, such as the soil surface, that restricts the movement of either water or solutes.
- **Pressure potential** (ψ_p) refers to the soil water potential attributable to any other factors besides gravity and solutes. Factors such as capillarity, adsorption, submergence, and air pressure can all influence the pressure potential.
  - **Matric potential** (ψ_m) is an important subcategory of pressure potential. Matric potential is the decrease in the water potential due to the interaction of water with the soil's solid matrix, primarily through capillarity and adsorption.

Soil water potential can be expressed in a variety of different ways, creating some opportunity for confusion. The three most common ways of expressing soil water potential are:

- **Energy per unit mass** (J kg^{-1}) is perhaps the most fundamental

and theoretically sound way to express soil water potential but is not as widely used the other two.

- **Energy per unit volume** (J m^{-3} or kPa) has the same units as pressure and is convenient for expressing pressure potential and osmotic potential. However, some uncertainty can arise with this form of expression since the density of water, and thus the relation of mass to volume, is slightly temperature dependent.
- **Energy per unit weight** (J N^{-1} or m or cm) is sometimes referred to as hydraulic head and, when the water flow velocity is negligibly small, is indicated by the height of the free water surface above or below a specified reference elevation. This way of expressing water potential is often convenient for solving soil water flow problems.

In this book we will primarily use energy per unit volume (kPa) and energy per unit weight (m or cm) to express soil water potential. Soil water potentials vary over several orders of magnitude. A soil in which the pore spaces are completely filled with water would have a matric potential of 0 kPa, while a surface soil exposed to dry atmospheric conditions for extended periods of time would have a matric potential far below -1500 kPa.

3.2.2 Soil Water Potential Measurement Techniques

Measurements of soil water potential or its components are used by researchers to understand water movement in the soil–plant–atmosphere continuum and by crop and turf managers to guide irrigation management. The oldest type of measuring device for this purpose is called a **tensiometer**, which consists of a ceramic cup connected to a water reservoir and a pressure gauge (Fig. 3-3). When the cup is embedded in the soil, water flows into or out of the

reservoir through the ceramic until the water potential inside the cup is equal to the water potential outside the cup.

Fig. 3-3. Tensiometer. Photo credit: Soil Moisture Equipment Corp. (link).

A more recent measurement technique for matric potential utilizes **heat dissipation sensors** which can provide matric potential measurements at potentials far below the tensiometer range. Heat dissipation sensors consist of a small temperature sensor and a heating element encased in the center of a ceramic plug (Fig. 3-4). When the sensor is embedded in soil, water moves into or out of the ceramic until the matric potential inside the ceramic equilibrates with the matric potential in the surrounding soil. A pulse of electric

current passed through the heater generates a temperature rise which is recorded using the temperature sensor.The movement of water into or out of the reservoir causes the pressure at the location of the pressure gauge to increase or decrease relative to atmospheric pressure. The gauge pressure, corrected for the height of the water column between the cup and the gauge, indicates the pressure potential or matric potential at the location of the cup. Tensiometers do not indicate osmotic potential because solutes can move freely through the ceramic. Tensiometers are typically limited to measuring pressure potentials greater than approximately -80 kPa, below which air passes through the ceramic and invalidates the pressure gauge reading.
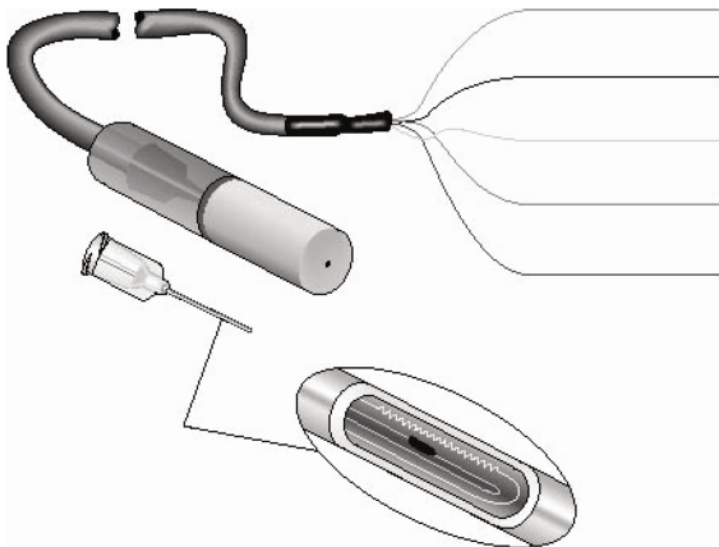
Fig. 3-4. Diagram of a heat dissipation sensor. Source: Campbell Scientific 229 Manual

Soil water potential can also be measured on soil samples in the

laboratory using an instrument such as a ***dewpoint potentiometer***. A small soil sample is placed inside the instrument in a sealed chamber containing a small mirror (Fig. 3-5). The water potential of the air in the space above the sample equilibrates with the soil water potential, and the instrument lowers the temperature of the mirror until finding the precise temperature at which dew forms on the mirror. Based on this dewpoint temperature, the water potential of the soil sample is determined. This instrument can measure water potentials over a large range but is even less sensitive in the wet range than are heat dissipation sensors. Current dewpoint potentiometers have an upper limit of approximately -100 kPa.The magnitude of the temperature rise is related to the absolute value of the matric potential. If the soil and ceramic are relatively dry, the temperature rise will be large, and the matric potential will be some large negative number. If the soil and ceramic are relatively wet, the temperature rise will be small, and the matric potential will be some small negative number. Heat dissipation sensors lose sensitivity when the ceramic is saturated and thus have an upper measurement limit of approximately -10 kPa.

Fig. 3-5. Dewpoint potentiometer. Source: METER Group (link).

3.3 Soil Water Retention

If you have ever taken a walk along a sandy beach, you probably observed that there is a place quite near the water's edge where the ground is dry enough and firm enough to easily walk on. In contrast, if you have walked along the edge of a lake or pond where the surrounding soil was fine-textured, you probably found that the ground near the water's edge was wet and muddy. The differences you experienced in those two cases can be partly explained by the differing capabilities of coarse- and fine-textured soils to retain, or store, water. These capabilities are described by a relationship called the soil water retention curve. The ***soil water retention curve*** is the relationship between soil water content and matric potential. Understanding this relationship is crucial to understanding processes such as soil water storage, water flow, and plant water uptake.

3.3.1   Features of Soil Water Retention Curves

The most fundamental concept to understand about soil water retention is that soil water content is positively related to soil matric potential. As soil water content decreases, matric potential also decreases, becoming more negative. When all the pores in a soil are filled with water, the soil is at its ***saturated water content*** (θ_s) and the matric potential is 0. Consider the water retention curve for the Rothamsted loam shown in Fig. 3-6. The intersection of the solid curve with the left-hand y-axis shows that for this soil θ_s is approximately 0.51 cm^3 cm^{-3}.

As we move to the right along the solid curve, we are moving toward more negative values of matric potential. The absolute value of matric potential, rather than matric potential itself, is plotted on

the x-axis in this figure, as is common for water retention curve plots. The absolute value of matric potential is sometimes called **suction**. Using the absolute value for matric potential allows us to use a logarithmic scale for matric potential to compensate for its large numerical range relative to that of soil water content.
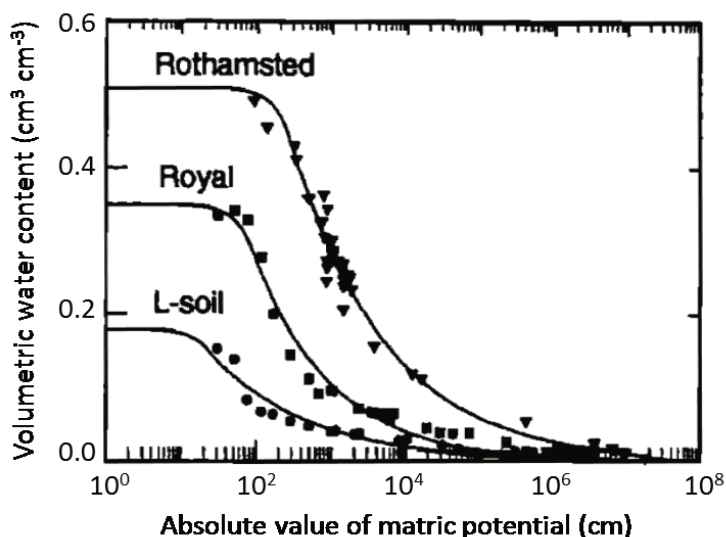
Fig. 3-6. Soil water retention curves for a sand (L-soil), a sandy loam (Royal), and a loam (Rothamsted). The symbols are measured data and the solid lines are models fit to the data. Adapted from Rossi and Nimmo (1994).

The water retention curve for the Rothamsted loam is flat between 10^0 cm (i.e. 1 cm) and approximately 10^2 cm (100 cm), then at lower matric potentials the curve bends downward. The highest matric potential at which air has displaced water in some of the pores of a previously saturated soil is called the **air-entry potential** (ψ_e). For this Rothamsted loam the air-entry potential was estimated to be -128 cm of water.

  As we follow the water retention curve toward the right from the

air-entry potential, we encounter a region where the decrease in water content is relatively large for each corresponding decrease in matric potential. There is a subtle *inflection point* approximately halfway down the descending limb of the water retention curve where the shape changes from concave to convex. The location of this inflection point may have some practical significance for soil management. The water content at this inflection point may be the optimum water content for tillage, resulting in the greatest proportion of small aggregates [17], and the slope of the curve at the inflection point may be a useful indicator of soil quality [18].

To the right of the inflection point, the steep portion of the curve tapers off into a relatively flat portion of the curve when the matric potential takes on large negative values. In this tail of the water retention curve, large decreases in matric potential are associated with only small decreases in soil water content.

3.3.2 Soil Properties Affecting Soil Water Retention

Another fundamental characteristic of soil water retention curves is that coarse-textured soils retain less water than fine textured soils at the same matric potential. Consider the substantial differences in the curves for the sand (L-soil), sandy loam (Royal), and loam (Rothamsted) textured soils in Fig. 3-6. The sand exhibits a much lower saturated water content than the loam, in this case 0.18 cm^3 cm^{-3} versus 0.51 cm^3 cm^{-3}. The sand also has a higher (less negative) air-entry potential than the loam, -32 cm versus -128 cm. The water retention for the medium-textured sandy loam soil is intermediate between those of the other two soils. Throughout the subsequent chapters, one common theme will be how these substantial differences in water retention between different soil textures dramatically influence water movement, plant growth, and related processes in both managed and natural ecosystems.

A secondary influence on soil water retention is the soil bulk density (Fig. 3-7). If you compare compacted and un-compacted samples of the same soil, the compacted soil will typically have a lower porosity, lower saturated water content, and lower air-entry potential. Sufficiently compacted soils can also have higher water contents for matric potentials below the air-entry potential than a similar un-compacted soil. This pattern is evident for the samples with the highest bulk density in Fig. 3-7 [19].
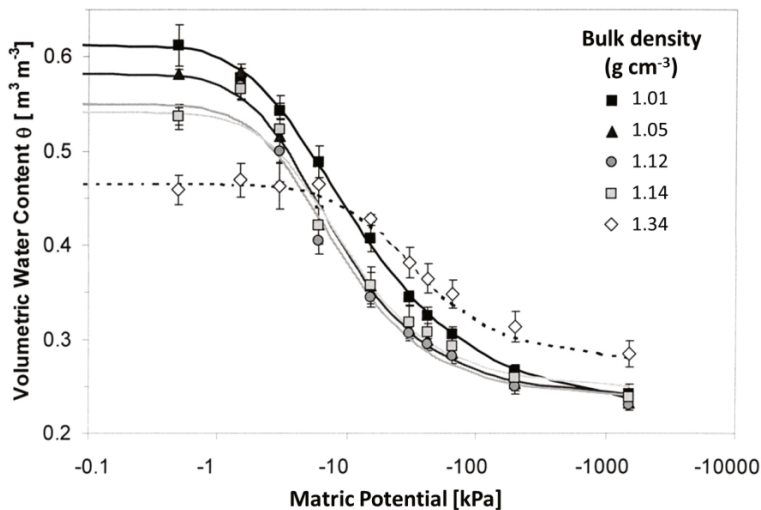
Fig. 3-7. Water retention curves for samples of a silt loam soil at different levels of compaction indicated by the bulk density values ranging from 1.01 to 1.34 g cm-3. Adapted from Stange and Horne (2005).

Advocates for conservation tillage, cover crops, soil quality, and, more recently, soil health have often stated that increasing soil organic matter improves soil water retention. However, the scientific evidence for this claim is somewhat unclear. While a number of studies have found that increasing organic matter increases soil water retention, a similar number of studies have

found no such effect [20]. One plausible hypothesis is that in some soils increasing organic matter results in decreased bulk density, leading indirectly to positive effects on water retention similar to those shown in Fig. 3-7.

3.3.3 Hysteresis in Soil Water Retention

The soil water retention curve can also be influenced by whether the soil is undergoing wetting (sorption) or drying (desorption). When the soil water retention curve differs between wetting and drying, that phenomenon is called **hysteresis**. This phenomenon has a number of important effects on soil water dynamics. For example, hysteresis in the water retention curve can increase the amount of water that is stored near the soil surface after an infiltration and drainage event [21]. Hysteresis can also slow the rate of solute leaching in soil under natural rainfall conditions with greater effects in coarse-textured than fine-textured soils [22]. In subsequent chapters, we will further consider the effects of hysteresis. For now, we will examine its nature and causes.

For a soil exhibiting hysteresis, the equilibrium water content associated with any particular matric potential will be lower for a wetting curve than for a drying curve (Fig. 3-8). The initial water content for the wetting or drying process also plays a role. Notice in Fig. 3-8 the clear difference in the drying curve for the silty clay loam soil when the drying process began from full saturation compared to when the drying process began at a lower water content indicated by the point labeled "B" [23].
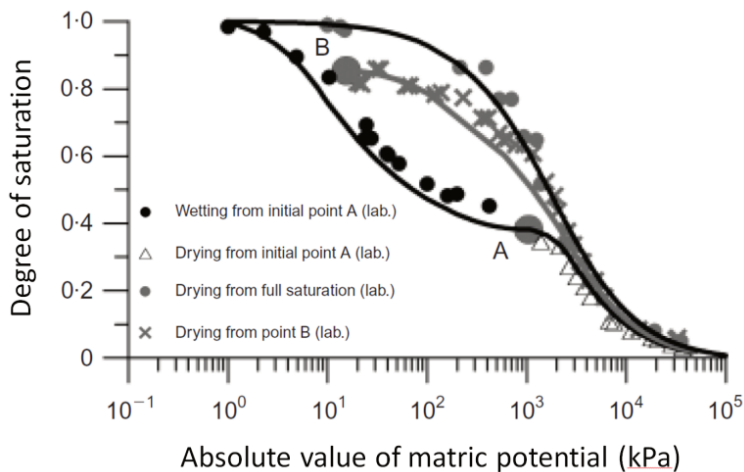
Fig. 3-8. Wetting and drying curves for an engineered silty clay loam soil exhibiting strong hysteresis for matric potentials between -1 and -1000 kPa. Degree of saturation is the volumetric water content divided by the saturated water content. Adapted from Tsiampousi et al. (2013).

Hysteresis in the soil water retention curve has multiple possible causes including: air entrapment, contact angle hysteresis, and the "ink bottle" effect. **Air-entrapment** occurs when a partially-drained soil is rewetted and small pockets of air become trapped in the interior pore spaces. This entrapped air cannot easily be removed, even if the soil is submerged underwater. As a result, higher water contents occur along the primary drainage curve from a fully saturated condition than those that occur during subsequent re-wetting (e.g. Fig. 3-8). Due to air-entrapment during re-wetting, the soil water content approaches a maximum value below the true saturated water content and this lower value is sometimes called the **satiated water content**. The image in Fig. 3-9 was generated by X-ray computed tomography and shows air-entrapment in the complex macropore network of a satiated soil column [24]. Soil chemical, physical, and biological processes can alter the amount

and distribution of entrapped air over time, so the impact of air-entrapment on soil water retention can change with each subsequent re-wetting cycle.
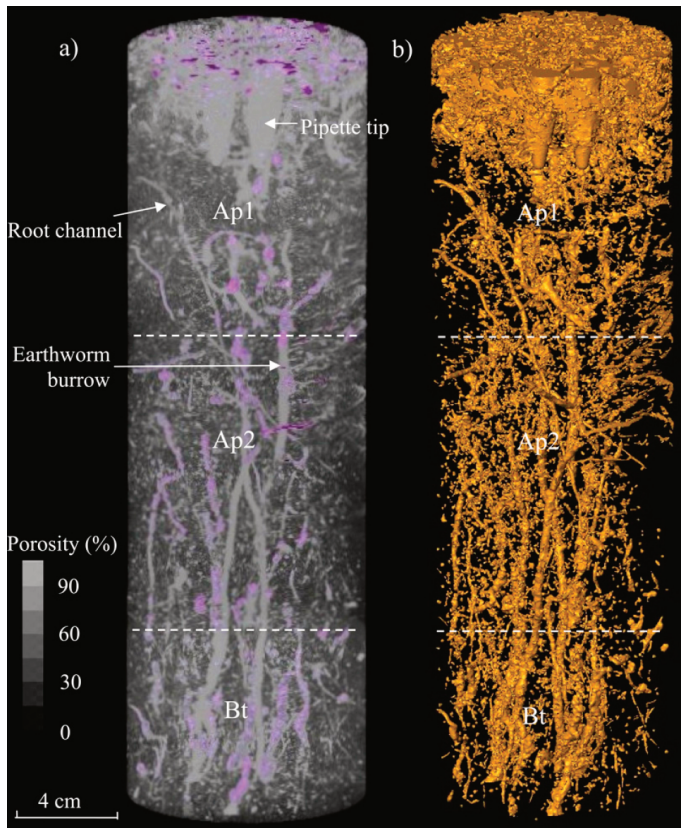
Fig. 3-9. Three dimensional visualizations of (a) porosity (gray) and entrapped air (pink) and (b) macropore network of a satiated 30-cm long column of a silt loam soil taken from the surface of cropland under no-till management. Adapted from Luo et al. (2008).

A second potential cause of hysteresis in the soil water retention curve is a phenomenon known as contact angle hysteresis. The **contact angle** is the angle at which a liquid-gas interface meets

a solid surface (Fig. 3-10). In our context, this means the angle at which the interface between the soil solution and the soil gas phase contacts the soil solids. Mineral soils often have contact angles <90° and are classified as **hydrophilic**, i.e. having affinity for water. Organic soils and mineral soils in which much of the surface area becomes covered with organic coatings can have contact angles >90°, making them **hydrophobic**, i.e. tending to repel water.

To visualize contact angle hysteresis and how it may affect soil water retention, a thought experiment may help. Imagine if we added a sufficiently small volume of liquid to the drop in Fig. 3-10a, the edge of the drop would not move but the contact angle would increase slightly. Likewise if we removed a sufficiently small amount of liquid, the contact angle would decrease slightly. Thus, contact angles for wetting and drying processes are different, i.e. contact angles exhibit hysteresis. The larger contact angles during wetting versus drying lead to higher (less negative) pressure potentials for the same water contents, consistent with Fig. 3-8.
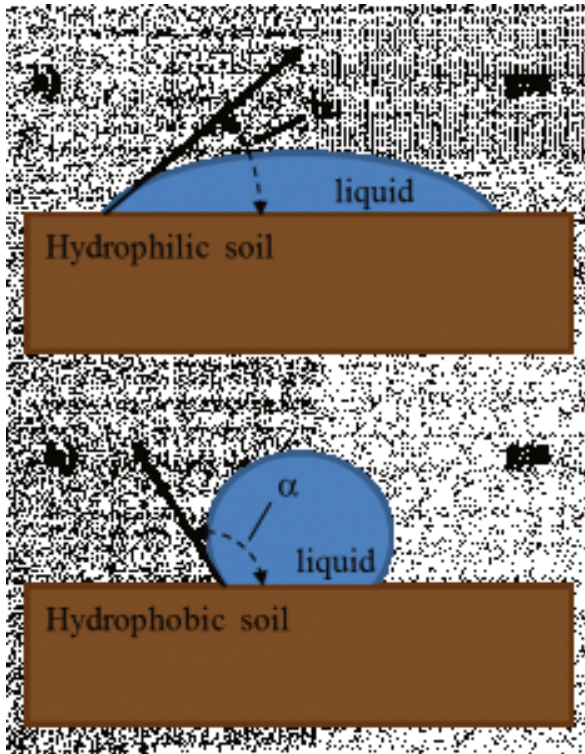
Fig. 3-10. Contact angles for hydrophilic and hydrophobic soils. The contact angle (a) is the angle from the solid surface to the liquid-gas interface, passing through the liquid.

A third potential cause for hysteresis is the **ink bottle effect**, which refers to the way in which drainage from a relatively large cavity, such as the body of an old-fashioned ink bottle, can be restricted if the fluid must drain through a relatively narrow opening, such as the neck of an inverted ink bottle. The analogy is somewhat helpful, but to better understand how this phenomenon influences soil water retention, we need to understand an important related phenomenon called capillary rise. **Capillary rise** is the rise of liquid against the force of gravity due to the upward force produced by

the attraction of the liquid molecules to a solid surface and to each other.

When you insert a small diameter tube, or capillary, into a fluid, such as water, the surface of the fluid inside the capillary may rise above that of the surrounding fluid, and the height (*h*) of this capillary rise is described by:

$$h = \frac{2\gamma cos\alpha}{\rho g r}$$

(Eq. 3-2)

where γ is the surface tension of the fluid (N m^{-1}), α is the contact angle of the liquid-gas interface on the wall of the tube, ρ is the fluid density (kg m^{-3}), g is the acceleration due to gravity (m s^{-2}), and r is the radius of the capillary (m). Thus, the smaller the radius of the capillary, the greater the height of the capillary rise. To better understand this equation, watch this video. The pressure potential just below the capillary meniscus is simply the negative of the capillary rise.
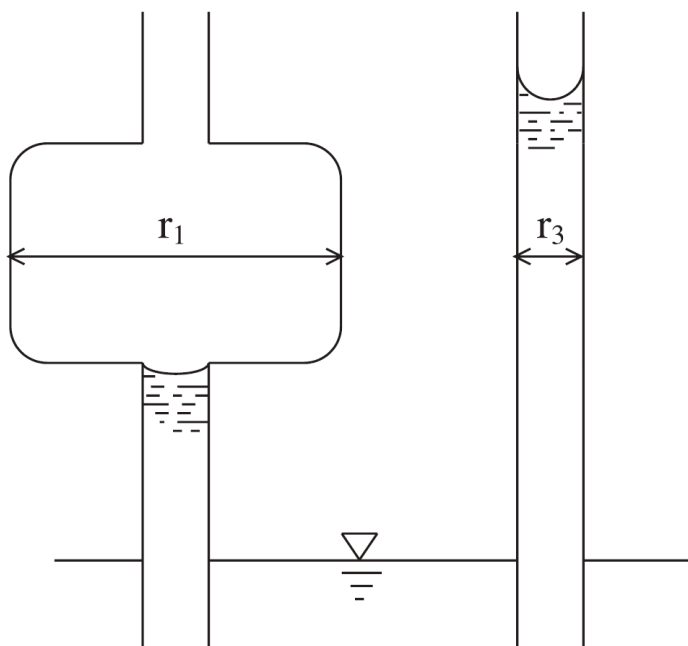
Fig. 3-11. Illustration of the "ink bottle effect" during wetting. Adapted from Ng and Pang (2000).

In Fig. 3-11, two capillary tubes have been inserted into water. The height of the resulting capillary rise was greater for the uniformly narrow tube on the right than for the non-uniform tube on the left. During this filling or wetting phase, capillary rise could only raise water to the bottom of the tube section with the enlarged diameter. If instead both tubes had drained from an initially water filled condition, then the enlarged section would have remained water-filled and height of water in both tubes would have been equal. Thus, for capillary tubes or soil pores with non-uniform radii, that non-uniformity can cause hysteresis in the water retention curve [25].

3.3.4 Measuring Soil Water Retention Curves

Because of the complexity of soil pore networks, we are currently unable to theoretically predict soil water retention curves from first principles, although progress has been made and is being made toward that goal [26, 27]. Until that goal is achieved, we will continue to determine soil water retention curves primarily by **empirical methods**, i.e. methods based on measurements and experience rather than theory or logical reasoning. Measurements of soil water retention are typically, but not always, performed in the laboratory with different methods being suitable for different portions of the possible range in soil matric potential. Near saturation, intact soil samples should be used because the soil structure and inter-aggregate pores can strongly influence water retention. At matric potentials below approximately −15 kPa, the effects of soil structure on water retention appear to be negligible [28] and smaller homogenized soil samples are typically used.
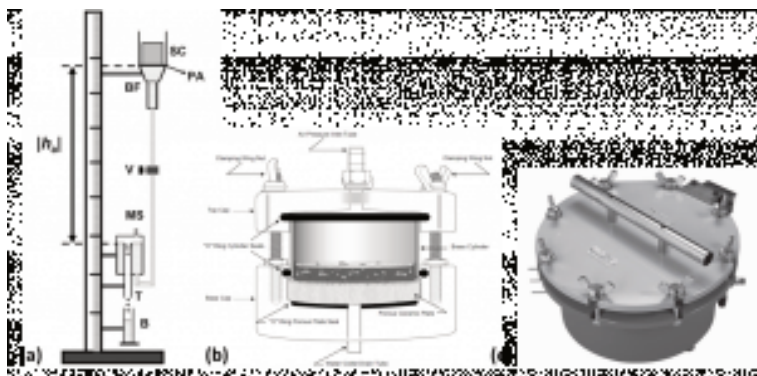
Fig. 3-12. Laboratory devices for measuring soil water retention include: hanging water column (a), Tempe cells (b), and pressure plate extractors (c). Image credits: (a) Nasta et al. (2011), (b) and (c) Soil Moisture Equipment Corp.

For matric potentials between 0 kPa and approximately −10 kPa, a simple **hanging water column** or tension table is often used to precisely control a sample's matric potential (Fig. 3-12a). When the

sample reaches equilibrium with the imposed matric potential, i.e. when water stops flowing, the water content of the sample can be determined by the change in the mass of the sample. For matric potentials between –10 and –100 kPa, small pressurized chambers often called *Tempe cells* work well, particularly for intact soil samples (Fig. 3-12b). A special porous ceramic plate at the bottom of the chamber, when saturated, allows water, but not air, to flow out of the chamber. The air pressure is increased to the absolute value of the desired matric potential, and once equilibrium is reached, the water content of the sample is determined based on the volume of water which flowed out of the sample or the change in mass of the sample. For matric potentials between –100 and –1500 kPa, specialized *pressure plates* in larger chambers have often been used (Fig. 3-12c). The principle of operation is similar to that of Tempe cells, but smaller samples of homogenized soil are used with each chamber housing multiple samples, and sometimes even multiple pressure plates. At these low matric potentials, true equilibrium may take many weeks or may never be reached, and a growing body of research suggests that data from pressure plate measurements may be unreliable at matric potentials below –100 kPa [29-31]. Dewpoint potentiometers (Fig. 3-5) offer one alternative measurement approach in this matric potential range.

3.3.5   Mathematical Functions for Soil Water Retention

Once we have measured soil water retention at several values of matric potential, we often need to fit a mathematical function to the measurements to allow calculation of water content for all other possible values of matric potential. One of the earliest widely-used water retention functions, developed by Brooks and Corey [32], is defined by:

$$\frac{\Theta - \Theta_r}{\Theta_s - \Theta_r} = (\frac{\Psi_c}{\Psi_m})^2 \, for \, \Psi_m < \Psi_c$$

(Eq. 3-3)

where θ_r is the ***residual water content***, which is conceptually the water content below which liquid water flow in the soil is no longer possible, ψ_e is the air-entry potential, and λ is a number related to the pore size distribution of the soil. Larger values of l indicate more uniformly-sized pores, while small values indicate a wide distribution of pore sizes are present.

A slightly simpler water retention function that is more convenient to use when performing calculations by hand was developed by Campbell [33] and is defined by:

$$\frac{\Theta}{\Theta} = (\frac{\Psi_e}{\Psi_m})^{\frac{1}{b}} \, for \, \Psi_m < \Psi_e$$

(Eq. 3-4)

where again *b* is a parameter related to the pore size distribution. The Campbell water retention function does not include a residual water content.

A more flexible and more widely-used water retention function was developed later by van Genuchten [34]. That function is defined by:

$$\frac{\Theta - \Theta_r}{\Theta_s - \Theta_r} = [1 + (-\alpha\Psi_m)^n]^{-m}$$

(Eq. 3-5)

where α is a parameter that is inversely related to the air-entry potential, n is a pore size distribution index similar to λ, and m is a parameter often defined by m = 1 – 1/n.

The most accurate way to estimate the parameters needed for these water retention functions is to obtain measurements of soil water retention across a broad range of matric potentials and then to adjust the parameters to achieve the best possible agreement with the measured values. Measured water retention curves for a

loamy sand and a silt loam soil are shown in Fig. 3-13 along with best-fits of the Brooks and Corey, Campbell, and van Genuchten water retention functions. All three functions fit the data reasonably, with the primary difference in this case being the sharp drop in water content at the air-entry potential predicted by the Brooks and Corey and the Campbell functions. The optimized parameters for each function are listed in Table 3-1 along with the root mean square error (RMSE), which is a measure of the error in the water content values estimated using the fitted function. For these two soils, all three functions fit the measured data well, but the van Genuchten function has the lowest RMSE.
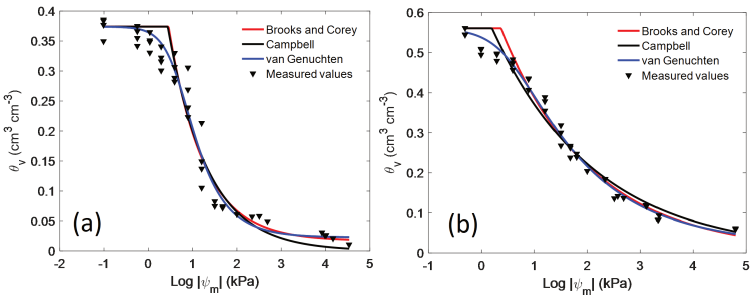
Fig. 3-13. Measured water retention curves for Tifton loamy sand (a) and Waukegan silt loam (b) plotted as black triangles along with solid lines showing the Brooks and Corey, Campbell, and van Genuchten water retention functions fit to the data.

If you do not have measurements of soil water retention for a particular soil, you can get a general idea of the shape of the water retention curve simply by knowing the soil textural class. Table 3-2 provides estimates of the parameters for the Brooks and Corey, van Genuchten, and Campbell water retention functions based on

Table 3-1. Best fit parameters for the Brooks and Corey, Campbell, and van Genuchten water retention functions for samples of Tifton sandy loam and

Soil	Function	Parameter				RMSE
		θ_s	θ_r	λ, b, or n	ψ_e or α	
		cm^3 cm^{-3}	cm^3 cm^{-3}	-	kPa or kPa^{-1}	cm^3 cm^{-3}
Tifton	Brooks & Corey	0.374	0.017	0.550	-2.87	0.031
sandy	Campbell	0.374	-	2.06	-2.75	0.032
loam	van Genuchten	0.374	0.023	1.74	0.211	0.025
Waukegan	Brooks & Corey	0.561	0.002	0.254	-2.34	0.030
silt loam	Campbell	0.561	-	4.49	-1.58	0.029
	van Genuchten	0.561	0.018	1.30	0.299	0.018

soil textural class alone. This table is one simple type of ***pedotransfer function***, a statistical tool for estimating unknown soil properties from known soil properties. The values in this table are suitable for educational purposes and general approximations of soil water retention behavior but not for many research or design purposes. For more reliable parameter estimates, you can use more complex and more accurate pedotransfer functions if you know additional soil properties such as percent sand, silt, and clay or bulk density or if you have one or more measurements of soil water retention available [35-37].

Table 3-2. Average parameters for the Brooks and Corey, van Genuchten, and Campbell soil water retention functions by USDA soil textural class. The residual water content (qr), saturated water content (qs), α, and n values were based on Schaap et al. (2001), the air-entry potential (ψ_e) and λ values were taken from Rawls et al. (1982), and the b values were taken from Rawls et al. (1992). Variables followed by a * are the back-transformed log mean for the textural class.

3.4 Problem Set

   1. A cylindrical soil sample of 3.85 cm diameter and 10.00 cm height weights 201.13 g. The sample is then oven dried at 105°C for 48 hours, reaching a final weight of 177.75 g. Assume a particle density of 2.65 g cm^{-3} and density of water 1.00 g cm^{-3}. Calculate the sample's:

     a. bulk density
     b. porosity
     c. gravimetric and volumetric water contents
     d. soil water storage in mm

2. Calculate the soil water storage in mm for a 60 cm deep soil profile in which the A horizon is 25 cm thick and has a volumetric water content of 0.25 cm^3 cm^{-3} and the B horizon is 35 cm thick and has a volumetric water content of 0.35 cm^3 cm^{-3}.

3. Determine the correct value for x in the simplified capillary rise equation below using Eq. 3-2 and assuming a contact angle of 20°. For water, assume a surface tension of 7.3 x 10^{-2} N m^{-1} (1 N m^{-1} = 1 kg s^{-2}) and density of 998 kg m^{-3}.

$$h = \frac{x}{r} cm^2$$

4. Use the result from the prior question to calculate the height of capillary rise in a 0.010 mm diameter capillary tube.

5. Use the Campbell water retention function and data from Table 3-2 to find the volumetric water content of a sand and a clay loam at -10. kPa and -1500. kPa.

3.5 References

1. Topp, C.G. and P.A. Ferré, *3.1 Water Content*, in *Methods of Soil Analysis: Part 4 Physical Methods*, J.H. Dane and C.G. Topp, Editors. 2002, Soil Science Society of America: Madison, Wisconsin. p. 417-546.

2. Entekhabi, D., et al., *The soil moisture active passive (SMAP) mission.* Proceedings of the IEEE, 2010. **98**(5): p. 704-716.

3. Kerr, Y.H., et al., *The SMOS mission: New tool for monitoring key elements of the global water cycle.* Proceedings of the IEEE, 2010. **98**(5): p. 666-687.

4. Larson, K.M., et al., *Use of GPS receivers as a soil moisture network for water cycle studies.* Geophys. Res. Lett., 2008. **35**(24): p. L24405.

5. Small, E.E., et al., *Validation of GPS-IR Soil Moisture Retrievals: Comparison of Different Algorithms to Remove Vegetation Effects.* IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016. **PP**(99): p. 1-12.

6. Dong, J., et al., *Calibration and validation of the COSMOS rover for surface soil moisture measurement.* Vadose Zone Journal, 2014.

7. Zreda, M., et al., *COSMOS: The COsmic-ray Soil Moisture Observing System.* Hydrology and Earth System Sciences, 2012. **16**(11): p. 4079.

8. Dong, J., et al., *Determining soil moisture and soil properties in vegetated areas by assimilating soil temperatures.* Water Resources Research, 2016. **52**(6): p. 4280-4300.

9. Sayde, C., et al., *Feasibility of soil moisture monitoring with heated fiber optics.* Water Resources Research, 2010. **46**(6): p. W06201.

10. Cosh, M.H., et al., *The Soil Moisture Active Passive Marena, Oklahoma, In Situ Sensor Testbed (SMAP-MOISST): Testbed Design and Evaluation of In Situ Sensors.* Vadose Zone Journal,

2016. **15**(4).

11. Robinson, D.A., et al., *Soil Moisture Measurement for Ecological and Hydrological Watershed-Scale Observatories: A Review.* Vadose Zone J, 2008. **7**(1): p. 358-389.

12. Topp, G.C., J.L. Davis, and A.P. Annan, *Electomagnetic determination of soil water content: Measurements in coaxial transmission lines.* Water Resources Research, 1980. **16**: p. 574-582.

13. Ochsner, T.E., et al., *State of the Art in Large-Scale Soil Moisture Monitoring.* Soil Science Society of America Journal, 2013. **77**(6): p. 1888-1919.

14. Krueger, E.S., et al., *Concurrent and antecedent soil moisture relate positively or negatively to probability of large wildfires depending on season.* International Journal of Wildland Fire, 2016. **25**(6): p. 657-668.

15. Rosenberg, E.A., A.W. Wood, and A.C. Steinemann, *Informing Hydrometric Network Design for Statistical Seasonal Streamflow Forecasts.* Journal of Hydrometeorology, 2013. **14**(5): p. 1587-1604.

16. Torres, G.M., R.P. Lollato, and T.E. Ochsner, *Comparison of Drought Probability Assessments Based on Atmospheric Water Deficit and Soil Water Deficit.* Agron. J., 2013. **105**(2): p. 428-436.

17. Dexter, A.R. and N.R.A. Bird, *Methods for predicting the optimum and the range of soil water contents for tillage based on the water retention curve.* Soil and Tillage Research, 2001. **57**(4): p. 203-212.

18. Dexter, A.R., *Soil physical quality: Part I. Theory, effects of soil texture, density, and organic matter, and effects on root growth.* Geoderma, 2004. **120**(3–4): p. 201-214.

19. Stange, C.F. and R. Horn, *Modeling the Soil Water Retention Curve for Conditions of Variable Porosity.* Vadose Zone Journal, 2005. **4**(3): p. 602-613.

20. Rawls, W.J., et al., *Effect of soil organic carbon on soil water retention.* Geoderma, 2003. **116**(1–2): p. 61-76.

21. Rubin, J., *Numerical Method for Analyzing Hysteresis-Affected,*

Post-Infiltration Redistribution of Soil Moisture1. Soil Science Society of America Journal, 1967. **31**(1): p. 13-20.

22. Persson, M. and A. Saifadeen, *Effects of hysteresis, rainfall dynamics, and temporal resolution of rainfall input data in solute transport modelling in uncropped soil.* Hydrological Sciences Journal, 2016. **61**(5): p. 982-990.

23. Tsiampousi, A., L. Zdravkovic, and D.M. Potts, *A three-dimensional hysteretic soil water retention curve.* Geotechnique, 2013. **63**(2): p. 155-164.

24. Luo, L., H. Lin, and P. Halleck, *Quantifying Soil Structure and Preferential Flow in Intact Soil Using X-ray Computed Tomography.* Soil Science Society of America Journal, 2008. **72**(4): p. 1058-1069.

25. Ng, C.W. and Y. Pang, *Experimental investigations of the soil-water characteristics of a volcanic soil.* Canadian Geotechnical Journal, 2000. **37**(6): p. 1252-1264.

26. Assouline, S., D. Tessier, and A. Bruand, *A conceptual model of the soil water retention curve.* Water Resources Research, 1998. **34**(2): p. 223-231.

27. Or, D. and M. Tuller, *Liquid retention and interfacial area in variably saturated porous media: Upscaling from single-pore to sample-scale model.* Water Resources Research, 1999. **35**(12): p. 3591-3605.

28. Nimmo, J.R., *Modeling Structural Influences on Soil Water Retention.* Soil Science Society of America Journal, 1997. **61**(3): p. 712-719.

29. Bittelli, M. and M. Flury, *Errors in Water Retention Curves Determined with Pressure Plates.* Soil Science Society of America Journal, 2009. **73**(5): p. 1453-1460.

30. Gee, G.W., et al., *The Influence of Hydraulic Nonequilibrium on Pressure Plate Data.* Vadose Zone Journal, 2002. **1**(1): p. 172-178.

31. Solone, R., et al., *Errors in water retention curves determined with pressure plates: Effects on the soil water balance.* Journal of Hydrology, 2012. **470–471**: p. 65-74.

32. Brooks, R. and A. Corey, *Hydraulic Properties of Porous Media.*

Hydrology Papers, Colorado State University, 1964(March).

33. Campbell, G.S., A *Simple Method for Determining Unsaturated Conductivity From Moisture Retention Data.* Soil Science, 1974. **117**(6): p. 311-314.

34. van Genuchten, M.T., A *Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils.* Soil Science Society of America Journal, 1980. **44**: p. 892-898.

35. Rawls, W., D. Brakensiek, and K. Saxtonn, *Estimation of soil water properties.* Transactions of the ASAE, 1982. **25**(5): p. 1316-1320.

36. Rawls, W.J., L.R. Ahuja, and D.L. Brakensiek, *Estimating soil hydraulic properties from soils data*, in *Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils*, M.T. van Genuchten, F.J. Leij, and L.J. Lund, Editors. 1992, Univ. of California, Riverside: Riverside, California. p. 329-340.

37. Schaap, M.G., F.J. Leij, and M.T. van Genuchten, ROSETTA: *a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions.* Journal of Hydrology, 2001. **251**(3-4): p. 163-176.

4. SOIL WATER FLOW

In the second chapter we considered the complex and hierarchical spatial organization of the soil solid phase across scales ranging from kilometers to micrometers. We briefly discussed the relevance of these patterns to issues ranging from climate change to agricultural management to pesticide transport. In the third chapter, we focused on two of the most important and dynamic descriptors of the soil's condition, soil water content and soil water potential, and the relationship between those two variables, the soil water retention curve. Now the main aim of this chapter is to build a foundation for accurately understanding soil water flow. But before we can do that, we first need to focus on some of the most fascinating, intricate, and life-giving physical features of soil, the soil pore spaces. We will begin this chapter by focusing on understanding soil pores and pore networks and their significance, and then we will spend the rest of the chapter learning the fundamental physical properties and processes that govern the flow of water in those pores spaces.

4.1 Pores and Pore Networks

Closely associated with the spatial patterns of the soil solid phase are the complementary spatial patterns of the soil pore network. The soil pores provide vital flow paths for water, oxygen, carbon dioxide, and nutrients without which life in the soil would be impossible. They also provide habitat for a host of living organisms in the soil. Soil scientists have long known that the soil pore networks are incredibly complex with different degrees of connectivity between pores giving rise to tortuous flow paths and with pore sizes varying by orders of magnitude even within the same soil. So before we can understand water flow in soil, we will need to understand soil pores and pore networks.

The largest soil pores are called macropores, and they are frequently major contributors to water flow and solute transport in soils (Fig. 4-1). The minimum size criteria for soil macropores is somewhat subjective, but with respect to water flow and transport of solutes, the available empirical evidence suggests macropores are those with equivalent diameters greater than approximately 0.3 mm [1]. Macropores can be created by organisms including soil-dwelling animals, plant roots, and human tillage practices, as well as through physical processes such as soil shrinkage (resulting in cracks) and internal erosion (resulting in natural soil pipes) [2].

Fig. 4-1. Macropores in a sub-soil core pulled from an alluvial soil along the south side of the Washita River near Chickasha, Oklahoma.

Biologically, macropores are often "hot-spots" where plant roots grow preferentially and soil organisms move more easily than in the bulk soil. As a result, macropore walls can be enriched in organic matter and clay content relative to the surrounding soil matrix (recall Fig. 2-16), and this enrichment in turn affects chemical and biological processes. For example, soil from macropore walls has shown greater sorption and more rapid mineralization of organic contaminants, greater sorption of trace metals, greater microbial activity, and more rapid N cycling [1].

It is not only the size but also the organization of the pores into pore networks which influences soil water flow. For example, large dead-end pores may contribute little to water movement relative to smaller pores which provide a continuous flow path. Our ability to "see through" intact soil to visualize 3-D pore networks is in its infancy, but researchers using technologies from the field of

medicine, such as X-ray computed tomography, are now able to provide intriguing images of pore networks that are changing our understanding of soil-water interactions. These images, like the ones in Fig. 4-2, reveal networks that can differ dramatically in their shape and connectivity and challenge our current mental and mathematical models of soil water flow and storage.

Leonardo Da Vinci has been quoted as saying, "We know more about the movement of celestial bodies than about the soil underfoot," and that is probably true with regards to our knowledge of macropores' effects on soil water flow. Our fundamental concepts and theories for soil water storage and flow were mostly developed in the first half of the 20th century by engineers and physicists, often working in a laboratory with homogenized soil, sand, or even glass beads, instead of intact soil in the field. Thus, the soil water fundamentals that were discovered then, and still form the core of our knowledge, largely ignore the existence and effects of macropores on soil water processes. One of the current grand challenges for soil physicists and hydrologists is to develop new, unifying concepts and quantitative models that effectively incorporate the now substantial body of information on soil macropores and their effects on water flow [3]. As we move now to consider established concepts of soil water flow that are fundamental to soil physics and hydrology, we need to keep in mind this grand challenge and the limitations of our existing concepts.
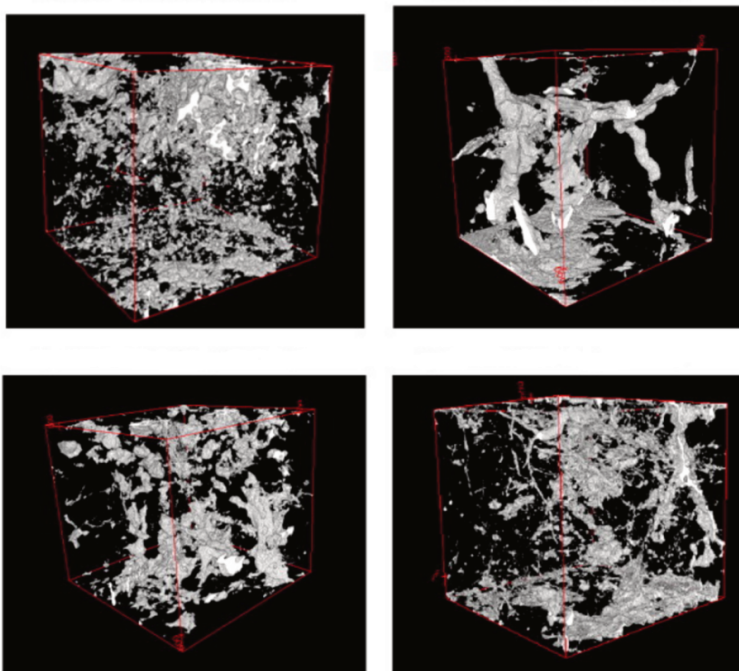
Fig. 4-2. Pore networks in four soil samples visualized using X-ray computed tomography (CT). The pores shown here are the air-filled pores larger than approximately 10 mm. The scanned regions are 36 mm on each side. Adapted from Munkholm et al. (2012).

4.2 Soil Water Potential for Systems at Equilibrium

Perhaps the most fundamental concept for understanding soil water flow is the fact that **differences in soil water potential drive soil water flow**. Intuitively, we might assume that water always flows downward through the soil, or perhaps we may assume that water always flows from wetter soil to drier soil. But reality can be surprising, and these intuitions can mislead us. In fact, water can and does flow upwards and horizontally in soil under certain conditions, and water can flow from drier soil to wetter soil under certain conditions. Therefore, to understand water movement in soil, we must supplement our intuitions with a solid understanding of the fact that differences in soil water potential ultimately drive soil water flow.

Fig. 4-3. Screenshot from "Soil water potentials in systems at equilibrium", an 8-minute video explaining how to analyze soil systems in hydraulic equilibrium. Click this link to view the video.

When soil water potential is the same throughout the soil, the soil is said to be in **_hydraulic equilibrium_**, and no water flow occurs. Although the total soil water potential is uniform throughout a soil that is at hydraulic equilibrium, the component potentials such as gravitational potential and pressure potential often vary throughout the soil. In fact, if we neglect any differences in osmotic potential, then for soil at hydraulic equilibrium, the variations in gravitational potential and pressure potential will perfectly offset one another. In this way the total potential is the same throughout the soil. Before you can solve soil water flow problems, you need to be able to correctly determine soil water potentials for systems in equilibrium. To do that you can use diagrams and tables like the ones shown in Fig. 4-3, as explained in the accompanying video. Please get a pencil and paper and watch the video now, taking time to create your own tables and fill in the blanks as you go.

4.3 Poiseuille's Law

When differences in soil water potential occur, **water flows from regions of higher potential to regions of lower potential**, unless those regions are separated by an impermeable layer. This knowledge alone is enough for us to determine the direction of soil water flow in any situation where we know or can measure the soil water potentials. But, if we want to estimate the rate of water flow through the soil, we need to know the relationship between the differences in water potential and the flow rate. To understand that relationship, let's begin by considering Poiseuille's Law for laminar flow through a tube:

$$\varrho = \frac{\pi r^4}{8\eta} \frac{\Delta p}{L}$$

(Eq. 4–1)

where Q is the volume of flow per unit time ($m^3\ s^{-1}$), r is the radius of the tube (m), Δp is the pressure difference from one end of the tube to the other (Pa), η is the dynamic viscosity of the fluid (Pa×s), and L is the length of the tube (m). Recall that soil water potentials can also be expressed as pressures, and you can see that the term Δp is analogous to the difference in water potential from one point in the soil to another. The term $\Delta p/L$ defines the **hydraulic gradient**, the ratio of the difference in hydraulic pressure to the distance over which that difference occurs. So, Poiseuille's Law tells us that the flow rate through a tube is proportional to the hydraulic gradient. If the hydraulic gradient decreases, the flow rate also decreases. We will see that the same principle holds true for soil water flow.

We can gain a second key insight from Poiseuille's Law. Often it is useful to consider not only the volumetric flow rate, Q, but also the **flux**, q, which is the volumetric flow rate per unit area ($m\ s^{-1}$). By inspecting Poiseuille's Law and recalling that the cross-sectional

area of a cylindrical tube is πr^2, we can see that the flux through a tube is given by:

$$q = \frac{r^2}{8\eta} \frac{\Delta p}{L}$$

<div align="right">**(Eq. 4-2)**</div>

The term r^2 shows that the magnitude of the flux for a particular hydraulic gradient depends strongly on the radius of the tube. If r is larger by a factor of 10 then q is larger by a factor of 100. We will see that a similar principle holds true for soil water flow; the water flux through the soil depends strongly on the size of the soil pores.

A third key insight we can gain from Poiseuille's Law is that properties of the fluid also influence the flow. The term η^{-1} shows that as the viscosity of the fluid increases the magnitude of the flux decreases. The viscosities of water and air, the two fluids most commonly found in soil, increase as temperature decreases. Thus, from Poiseuille's Law, we can infer that flow rates for these fluids through soil will decrease as the temperature decreases, and they do.

Although Poiseuille's Law, published in 1841, allows us to gain some helpful insights about soil water flow, it is of little practical value for solving soil water flow problems. Soil is not a smooth straight tube, nor is it a bundle of smooth straight tubes. We usually cannot identify any meaningful "radius" for the soil pore network which would allow us to directly apply Poiseuille's Law to the soil. We will need another approach.

4.4 Darcy's Law

In 1856, a French engineer named Henry Darcy (Fig. 4-4) working on water filtration systems proved that water flow through sand beds was proportional to the hydraulic gradient, the same relationship Poiseuille had discovered for laminar flow through tubes [4]. The proportionality constant, K, is called the **hydraulic conductivity** and is a measure of a material's ability to transmit water. Darcy's discovery has come to be called Darcy's Law and can be written as:

$$q = K_s \frac{\Delta \Psi_t}{L}$$

<div align="right">

(Eq. 4-3)

</div>

where $\Delta\psi_t$ is the difference in total water potential between two points in a saturated porous media separated by distance, L. The symbol K_s serves to clarify that we are referring to the **saturated hydraulic conductivity** of the soil, which differs dramatically from the hydraulic conductivity of unsaturated soil as we will soon see. Darcy noted that the hydraulic conductivity depended on the permeability of the porous media. By comparing Darcy's Law with Poiseuille's Law, we can also see that the hydraulic conductivity depends on the viscosity of the fluid, which itself with vary with the temperature of the fluid. Although the experiments leading to Darcy's law used only sand, the relationship has proven to accurately describe flow in a wide array of soil types and other porous media. It has become one of the most important relationships in soil physics, hydrology, and hydrogeology. Please get a pencil and paper and watch the following two videos to learn how to use Darcy's Law to estimate water flow through saturated soil columns. The first video shows an example of flow through a horizontal soil column, and the second video shows an example of flow through a vertical soil column. Pause the videos as necessary and work the problems for yourself to increase your understanding.

Fig. 4-4. Portrait of Henry Darcy by F. Perrodin from the Collection of the Bibliothéque Municipale de Dijon. Public domain.

Where Poiseuille's Law required specific information about the flow geometry, i.e. the radius of the tube, Darcy's Law requires only an empirically determined constant, the hydraulic conductivity. This gives the Darcy's Law the flexibility to apply to virtually any porous media, as long as the flow is laminar. But, this flexibility comes at

a cost; the hydraulic conductivity can differ by several orders of magnitude from one soil type to another. The hydraulic conductivity can also differ by an order of magnitude or more from one sample to another within the same soil type. To better understand the causes of this enormous variability, let's consider some of the primary factors which influence hydraulic conductivity and soil water flow rates. We will here be focusing, as Darcy did, on the hydraulic conductivity of saturated soil, K_s.

4.5 Factors Affecting Saturated Hydraulic Conductivity

As you may have expected, soil texture strongly influences saturated hydraulic conductivity. Soils dominated by large sand particles tend to have relatively large pore spaces and thus large values of saturated hydraulic conductivity. Soils dominated by small clay particles tend to have relatively small pores spaces and small values of saturated hydraulic conductivity. For example, using intact soil cores, Reynolds et al. measured saturated hydraulic conductivity values of 29 cm h^{-1} for a sand, 4.1 cm h^{-1} for a loam, and 0.091 cm $^{-1}$ for a clay loam in Canada [5]. The saturated hydraulic conductivity of the sand was more than 300 times larger than that of the clay loam.
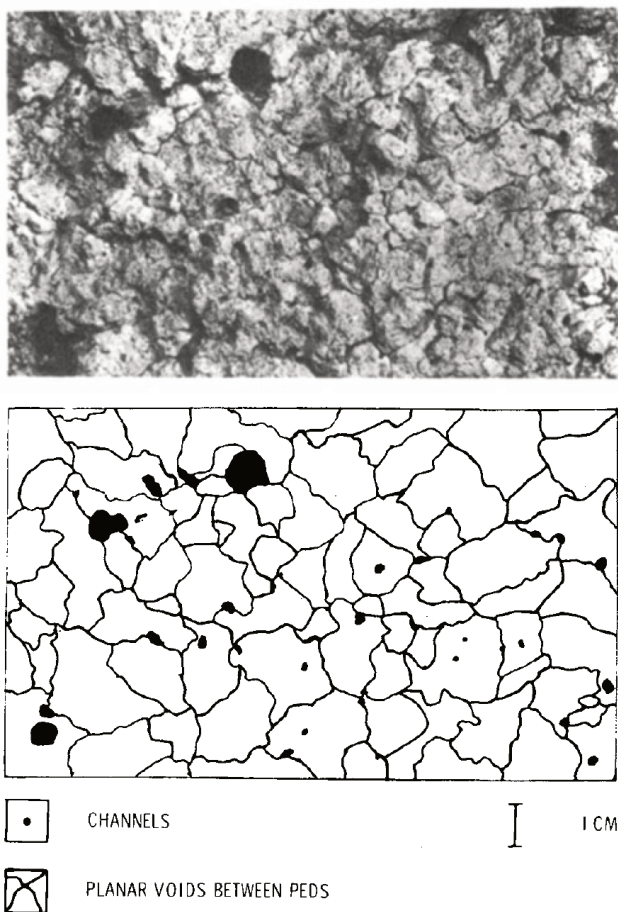
Fig. 4 5. Photograph of a horizontal cross-section through the Bt horizon of a Batavia silt loam soil in Wisconsin (upper panel). Tracing of structural features highlighting macropores and cracks between aggregates (lower panel). Reproduced from Anderson and Bouma (1973).

The presence, size, and continuity of macropores can also strongly influence saturated hydraulic conductivity. Poiseuille's Law indicates that the flux through a tube increases with the square

of the radius, and water flux through soil is also quite sensitive to the presence of large pores, even if they are few in number. Yet in the soil many other features of the pores are involved such as the connectivity, the internal roughness, and the **tortuosity**, which is a measure of the extent of twists and turns taken by the pores. Clayey soils with large numbers of well-connected macropores generated by living organisms in the soil, e.g. earthworms, can have saturated hydraulic conductivity values greater than those of coarse-textured soils which lack macropores [6].

Likewise, the type and degree of soil structure also affect saturated hydraulic conductivity. Strongly developed, fine blocky soil structure contributes to high values of saturated hydraulic conductivity, while massive, featureless soil structure often indicates compaction and low values of saturated hydraulic conductivity [6]. By carefully measuring soil structural features such as the length, width, and number of inter-aggregate pores (Fig. 4-5) and applying an appropriate model, one can reasonably predict the saturated hydraulic conductivity of a soil horizon [7]. In fact, some US states use visual inspection of the soil profile by trained soil scientists as a primary factor in determining the suitability of a location for the drain field of an on-site waste water treatment system, i.e. a septic system. The profile inspection allows the scientist to reasonably estimate the rate of water flow which the soil can sustain.

4.5.1   Chemical Dispersion and Flocculation

In addition to these physical properties, chemical properties of the soil and the solution flowing through the soil can also impact the saturated hydraulic conductivity. These chemical effects arise when soil and solution characteristics promote swelling and chemical dispersion of clay present in the soil. **Chemical dispersion** here is the process in which soil particles, which previously were held together

in close contact within soil aggregates, respond to a changed chemical environment by expanding and separating from one another, breaking down the soil aggregates. Swelling and dispersion can reduce the saturated hydraulic conductivity of a soil by a factor of 100 or more [8]. Swelling and dispersion can be promoted by any of the following conditions: irrigation with sodic water [8]; a high content of 2:1 clays, particularly montmorillonite [9]; low electrical conductivity of the flowing solution [10]; or high exchangeable sodium percentage (ESP) in the soil. The ESP is defined based on the amount of exchangeable sodium in a soil divided by the sum of the exchangeable calcium, magnesium, potassium, and sodium. For example, leaching a column of sandy loam soil having an initial ESP of 10% with distilled water caused a 90% reduction in the K_S value (Fig. 4-6).

Chemical dispersion of soil can sometimes be reversed, and the reverse process is called flocculation. **Flocculation** is the process in which dispersed soil particles come together, often due to a change in the chemical environment. A high proportion of polyvalent cations, such as Ca^{2+}, Mg^{2+}, and Al^{3+}, promotes flocculation, while a high proportion of monovalent cations, particularly Na^+, promotes chemical dispersion. Gypsum ($CaSO_4$) has been successfully used to remediate chemical dispersion in sodic and saline-sodic soils because it provides Ca^{2+} to displace Na^+ on the soil's cation exchange sites [11]. Sulfuric acid (H_2SO_4) is another good remediation option for chemical dispersion in calcareous sodic soils because the acid dissolves calcium carbonate ($CaCO_3$) present in the soil, and the Ca^{2+} released displaces Na^+ [11].
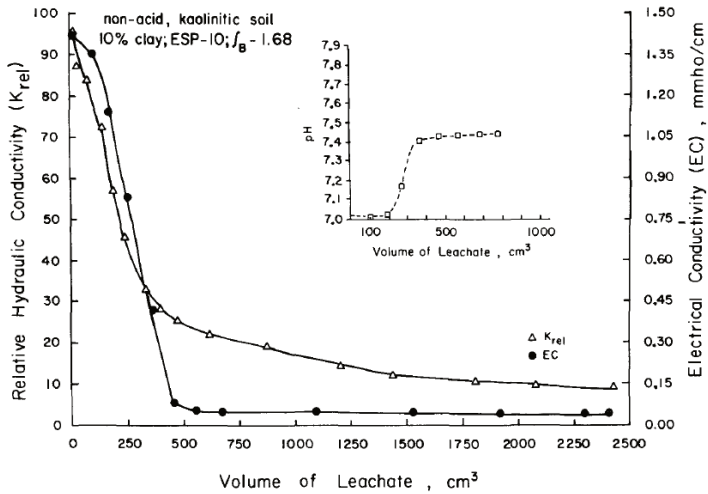
Fig. 4-6. *Relative hydraulic conductivity (Krel) and effluent electrical conductivity (EC) and pH changes produced by leaching a column of Fallbrook sandy loam with distilled water. The soil had 10% of primarily kaolinitic clay, an initial exchangeable sodium percentage (ESP) of 10 and a bulk density of 1.68 g cm^{-3}. Reproduced from Frenkel et al. (1978).*

4.6 Darcy's Law for Layered Soil

If saturated hydraulic conductivity can vary by orders of magnitude, then what does that mean for saturated flow through soil profiles, which consist of horizons with differing properties? Watch this video to learn how Darcy's Law can be rearranged into a convenient form for solving saturated water flow problems for layered soil. For a soil with two distinct layers, the Darcy's Law can be written as:

$$q = \frac{\Delta \Psi_t}{R_{h1} + R_{h2}}$$

where R_{h1} and R_{h2} are the hydraulic resistances for layers 1 and 2, respectively. The **hydraulic resistance** of a soil layer is simply the thickness of the layer (L) divided by the hydraulic conductivity of the layer, R_h = L/K. For a soil consisting of more than two layers, additional hydraulic resistances can be added to the denominator of the equation above for each layer. This video shows an example of applying Darcy's Law for layered soil to calculate flow through a soil profile with two distinctly different layers. Please take the time to watch the video as many times as necessary, following along with pencil and paper until you are confident you can solve this type of flow problem.

Darcy's Law for layered soil allows us to estimate water flow rates for layered soils, which we are likely to encounter in the field. However, both forms of Darcy's Law only apply to saturated water flow. Many times in the field we need to understand or estimate water flow rates when the soil is unsaturated. For example, infiltration into, redistribution through, and drainage from soil profiles all generally involve unsaturated soil. It would be 50 years

after Darcy's landmark paper before someone developed a relationship to predict water flow in unsaturated soil.

4.7 Buckingham-Darcy Law

In 1902, a physicist named Edgar Buckingham was hired by the US Department of Agriculture to work in its newly formed Bureau of Soils [12]. He stayed there for only four years, but his studies during that time led to a conceptual breakthrough that would change the course of soil physics and hydrology. Drawing on earlier work by John Maxwell and Lyman Briggs, Buckingham reasoned that in unsaturated soil the water was attracted to and held by the surfaces of the soil solids by what he called "capillary potential", i.e. matric potential [13]. He saw that spatial gradients in this potential would act as a force to drive soil water flow and that the resulting flow would be proportional to and in the opposite direction of that gradient. Buckingham also correctly predicted that in unsaturated soil the "capillary conductance", i.e. the hydraulic conductivity, would be largely dependent on the soil water content. And, Buckingham measured the variation of the equilibrium soil water content with height above a water table in long laboratory soil columns, resulting in the first published soil water retention curves relating water content and "capillary potential".

Because of Buckingham's careful experimental work and perceptive theoretical insights, subsequent soil physicists were able to formulate a single equation which describes both saturated and unsaturated water flow in soil and other porous media:

$$q = -K(\Theta)\frac{d(\Psi_p + \Psi_g)}{dz}$$

<div align="right">(Eq. 4-5)</div>

where $K(\theta)$ indicates that the hydraulic conductivity is a function of the soil volumetric water content and $d(\psi_p + \psi_g)/dz$ indicates the gradient of the total water potential in the z direction. Osmotic potential can also be included with the other potentials if necessary. This equation can be viewed as a generalization of Darcy's Law and can be referred to as the ***Buckingham-Darcy Law***.

To understand unsaturated water flow in soil, we need to appreciate at least three key differences between Darcy's Law and the Buckingham-Darcy Law. First, we need to understand that the hydraulic conductivity depends strongly on the soil water content in unsaturated soil and that water content can vary in space and time. For example, the hydraulic conductivity of Grenoble sand is near 15 cm h^{-1} when the water content is 0.30 cm^3 cm^{-3}, but the hydraulic conductivity drops more than one order of magnitude, to approximately 1 cm h^{-1} when the water content decreases to 0.20 cm^3 cm^{-3} (Fig. 4-7, [14]). The cause of this dramatic sensitivity to water content is linked to what we learned previously about the soil water retention curve, capillary rise, and Poiseuille's Law.
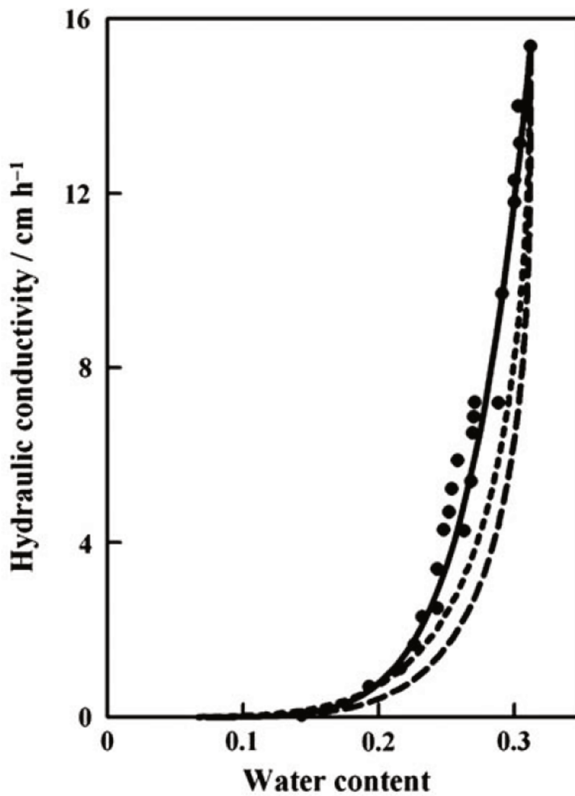
Fig. 4-7. Hydraulic conductivity as a function of volumetric water content for Grenoble sand. The circles are the measured data and the lines are estimates from three models. Reproduced from Touma (2009).

From the soil water retention curve, we know that soil water content is positively related to soil matric potential, so if the soil water content decreases from 0.30 to 0.20 cm^3 cm^{-3}, a corresponding decrease in matric potential must occur. From the capillary rise equation, we can infer that as matric potential decreases (analogous to an increase in capillary rise), water is held in capillary spaces with smaller radii, r. And from Poiseuille's Law,

we know that flow through a capillary depends on r^2. So if the capillary radius is reduced by a factor of 4, then the flow would be reduced by a factor of 16. Now, soils are not capillary tubes, but this line of reasoning helps us to understand and explain why hydraulic conductivity depends so strongly on soil water content.

The second key difference between Darcy's Law and the Buckingham-Darcy Law is that, for the latter, the gradient driving the flow depends on the soil water content. The gradient term in the Buckingham-Darcy Law includes ψ_p, the pressure potential. For unsaturated soil the pressure potential term includes the matric potential, ψ_m, which is related to the soil water content by the water retention curve. Thus, within an otherwise homogeneous soil, gradients in water content indicate gradients in matric potential which drive soil water flow from areas of higher water content to areas of lower water content.

A third key difference between Darcy's Law and the Buckingham-Darcy Law is a consequence of the first two differences highlighted. Darcy's Law can be applied directly to determine the flux through a layer of saturated soil of any desired thickness, so long as the hydraulic conductivity and the hydraulic gradient are defined for that layer. But, the Buckingham-Darcy Law as written above cannot be applied directly to determine the flux through a layer of unsaturated soil when the water content varies throughout the layer. In those situations, the hydraulic conductivity and the hydraulic gradient can vary by orders of magnitude within the layer, so no single value for those variables can be used to represent the layer. That is why the Buckingham-Darcy law is written as a ***differential equation*** with the term $d(\psi_p + \psi_g)/dz$ indicating that the gradient is defined for a layer of infinitesimal thickness. The corresponding hydraulic conductivity must also be evaluated for the water content of that infinitesimal layer, and the product of the gradient and the conductivity gives the flux for that vanishingly thin layer.

To calculate the flux across a real soil layer of finite thickness, the Buckingham-Darcy Law must be integrated across that layer.

We will encounter an example of this approach later when we study evaporation from an underground water table. However, such integration is possible only for some relatively simple and specific flow problems. In other cases, it is possible to apply some simplifying assumptions to the Buckingham-Darcy Law, such as assuming a unit hydraulic gradient, resulting in an algebraic expression to calculate the flux at a particular depth in the soil. We will see an example of this approach when we study soil water redistribution and deep drainage.

4.8 Models for Soil Hydraulic Conductivity

We sometimes have measurements of soil hydraulic conductivity at saturation and perhaps at one or two water contents below saturation, but we often need a mathematical function to allow calculation of hydraulic conductivity for all other values of water content. For this reason, soil hydraulic conductivity functions have been developed corresponding to each of the soil water retention functions presented in Chapter 3. The hydraulic conductivity function of Brooks and Corey [15], is defined by:

$$\frac{K(\Theta)}{K_s} = \left(\frac{\Theta - \Theta_r}{\Theta_s - \Theta_r} \right)^{\frac{2+3\lambda}{\lambda}}$$

(Eq. 4-6)

where $K(\theta)$ is the hydraulic conductivity as a function of volumetric water content, K_s is the saturated hydraulic conductivity, and λ is the same pore-size distribution index used in the Brooks and Corey water retention curve. Larger values of λ indicate more uniformly-sized pores, while small values indicate a wide distribution of pore sizes are present.

The hydraulic conductivity function corresponding to the water retention model of Campbell [16] is defined by:

$$\frac{K(\Theta)}{K_s} = \left(\frac{\Theta}{\Theta_s} \right)^{2b+3}$$

(Eq. 4-7)

where again b is a parameter related to the pore size distribution. The most commonly-used hydraulic conductivity function

corresponding to the water retention model of van Genuchten [17] is defined by:

$$\frac{K(\Theta)}{K_s} = \left(\frac{\Theta - \Theta_r}{\Theta_s - \Theta_r}\right)^L \left\{1 - \left[1 - \left(\frac{\Theta - \Theta_r}{\Theta_s - \Theta_r}\right)^{n/(n+1)}\right]^m\right\}^2$$

(Eq. 4-8)

where n is a pore-size distribution index similar to λ, and m is a parameter defined in this case by $m = 1 - 1/n$.

4.9 Conclusion

In this chapter, we have considered the complexity of soil pores and pore networks and how they control and complicate the processes of soil water flow. We have learned how to make a simple table to determine the gravitational, pressure, and total water potentials for soil-water systems at equilibrium and during steady, saturated flow. We have studied, compared, and contrasted three foundational laws relevant to water flow in soil: Poiseuille's Law, Darcy's Law, and the Buckingham-Darcy Law. And, we have learned about soil hydraulic conductivity; its role in determining soil water flow rates, the enormous variability which occurs in hydraulic conductivity, and the soil and water factors which contribute to that variability. Equipped with understanding of these fundamental concepts, we are now ready to begin studying one-by-one the processes of the soil water balance, that dynamic ebb and flow which sustains life on Earth. As we will see in the next chapter, our study must begin before the water ever reaches the soil.

4.10 Problem Set

1. A soil profile is in hydraulic equilibrium with a water table located at a depth of 275 cm below the surface. Draw a sketch representing that soil profile, label the soil surface as point "A" and the water table as point "B". Create a table showing the gravitational head, pressure head, and total head at points A and B.

2. An initially saturated soil sample is brought into hydraulic contact with a thin porous plate connected to a hanging water column (see Fig. 3-12a) with a vertical length of 55 cm.

> a. When the soil reaches hydraulic equilibrium, what is the pressure head at the base of the soil sample?
>
> b. In one complete sentence explain why this pressure head occurs.

3. A cylindrical soil column of 100. cm^2 cross-sectional area and 50.-cm height is filled with homogeneous soil and saturated, and 10. cm of water is kept ponded on the surface. The soil column is open to the atmosphere and freely draining at the bottom. The steady-state volumetric flow rate through the soil column is 1000. cm^3 h^{-1}.

> a. Draw a sketch of this soil column.
>
> b. Create a table to determine the difference in hydraulic head across the column.
>
> c. Convert the volumetric flow rate to water flux (cm h^{-1}).
>
> d. Calculate the saturated hydraulic conductivity of the soil.

4. A 1.0-mm diameter tube is pushed through the column described in problem 3 and hollowed out. Steady state flow is established, with water flowing through the soil in accordance with Darcy's Law and through the tube in accordance with Poiseuille's Law. Assume the viscosity of water is 1.0 x 10^{-3} kg m^{-1} s^{-1}.

> a. Calculate the volumetric flow rate through the tube (cm3 h-1).

b. Calculate the flux for the combined column-tube system in cm h-1. You can assume that the tube takes up a negligible portion of the cross-sectional area of the column, so the flow through the soil matrix itself is unchanged.

c. Calculate the effective saturated hydraulic conductivity for the combined column/tube system.

d. Write one sentence explaining the practical significance of this example.

Hint for #4: Convert your water potential difference from head units (e.g. cm) to pressure units (e.g. Pascals [Pa] which is the same as $kg\ m^{-1}\ s^{-2}$) for use in Poiseuille's Law. To do that, first express the water potential difference in meters and use the fact that $P = \rho_w\ g\ H$ where ρ_w is the density of water ($1000\ kg\ m^{-3}$) and g is acceleration due to gravity ($9.81\ m\ s^{-2}$).

5. A saturated soil column contains two soil layers, each 10. cm thick, with sand having $K_s = 10.\ cm\ h^{-1}$ underneath loam having $K_s = 5.0\ cm\ h^{-1}$, and 10. cm of water is kept ponded on the surface. The bottom of the soil column is open to the atmosphere and drains freely.

a. Draw a sketch of this soil column.
b. Create a table to determine the difference in hydraulic head across the column.
c. Calculate the hydraulic resistance for each layer.
d. Calculate the flux of water through the column using Darcy's Law for layered soil.
e. Calculate the pressure head at the sand-loam interface.

4.11 References

1. Jarvis, N.J., *A review of non-equilibrium water flow and solute transport in soil macropores: principles, controlling factors and consequences for water quality.* European Journal of Soil Science, 2007. **58**(3): p. 523-546.

2. Beven, K. and P. Germann, *Macropores and water flow in soils.* Water resources research, 1982. **18**(5): p. 1311-1325.

3. Beven, K. and P. Germann, *Macropores and water flow in soils revisited.* Water Resources Research, 2013. **49**(6): p. 3071-3092.

4. Brown, G.O., *Henry Darcy and the making of a law.* Water Resources Research, 2002. **38**(7): p. 11-1-11-12.

5. Reynolds, W.D., et al., *Comparison of Tension Infiltrometer, Pressure Infiltrometer, and Soil Core Estimates of Saturated Hydraulic Conductivity.* Soil Science Society of America Journal, 2000. **64**(2): p. 478-484.

6. McKeague, J., C. Wang, and G. Topp, *Estimating saturated hydraulic conductivity from soil morphology.* Soil Science Society of America Journal, 1982. **46**(6): p. 1239-1244.

7. Anderson, J.L. and J. Bouma, *Relationships Between Saturated Hydraulic Conductivity and Morphometric Data of an Argillic Horizon1.* Soil Science Society of America Journal, 1973. **37**(3): p. 408-413.

8. Frenkel, H., J.O. Goertzen, and J.D. Rhoades, *Effects of Clay Type and Content, Exchangeable Sodium Percentage, and Electrolyte Concentration on Clay Dispersion and Soil Hydraulic Conductivity1.* Soil Science Society of America Journal, 1978. **42**(1): p. 32-39.

9. McNeal, B.L. and N.T. Coleman, *Effect of Solution Composition on Soil Hydraulic Conductivity1.* Soil Science Society of America Journal, 1966. **30**(3): p. 308-312.

10. Zhang, X.C. and L.D. Norton, *Effect of exchangeable Mg on saturated hydraulic conductivity, disaggregation and clay*

dispersion of disturbed soils. Journal of Hydrology, 2002. **260**(1–4): p. 194-205.

11. Qadir, M., et al., *Amelioration strategies for sodic soils: a review.* Land Degradation & Development, 2001. **12**(4): p. 357-386.

12. Narasimhan, T.N., *Buckingham*, 1907. Vadose Zone Journal, 2005. **4**(2): p. 434-441.

13. Buckingham, E., *Studies on the movement of soil moisture.* 1907.

14. Touma, J., *Comparison of the soil hydraulic conductivity predicted from its water retention expressed by the equation of Van Genuchten and different capillary models.* European Journal of Soil Science, 2009. **60**(4): p. 671-680.

15. Brooks, R. and A. Corey, *Hydraulic Properties of Porous Media.* Hydrology Papers, Colorado State University, 1964(March).

16. Campbell, G.S., *A Simple Method for Determining Unsaturated Conductivity From Moisture Retention Data.* Soil Science, 1974. **117**(6): p. 311-314.

17. van Genuchten, M.T., *A Closed-form Equation for Predicting the Hydraulic Conductivity of Unsaturated Soils.* Soil Science Society of America Journal, 1980. **44**: p. 892-898.

5. WATER INPUTS
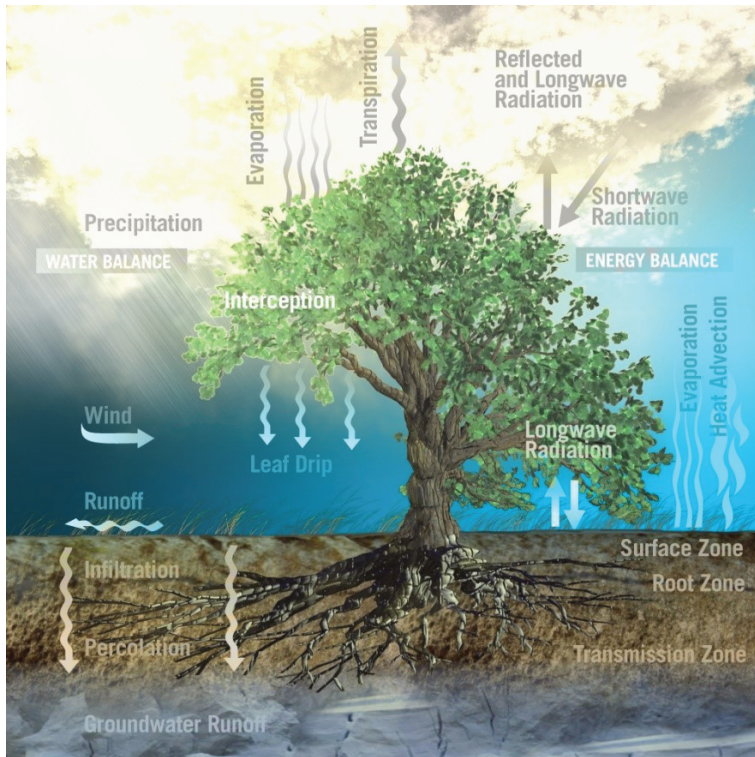
Fig. 5-1. The processes of the soil water balance (left side) and the land surface energy balance (right side). Source: European Space Agency (link).

The diagram in Fig. 5-1 will help to guide our systematic study of the processes involved in the soil water balance and subsequently our examination of land surface energy balance process. We will begin with consideration of the process of precipitation, the primary water input to the soil water balance in most environments. We will also briefly consider water inputs through irrigation. From there

we will follow the water as it moves through the soil water balance processes, considering in turn:

- interception of water by plant canopies and residue,
- impact of water drops on the soil surface,
- infiltration of water into the soil,
- the processes of runoff and erosion by water,
- then redistribution of water within the soil profile,
- drainage of water from the bottom of the soil profile,
- the associated processes of groundwater pollution and soil salinization,
- then evaporation of water from bare soil surfaces and erosion by wind,
- then root water uptake and transpiration by plants.

When we have completed this overview of the soil water balance processes, we will then consider the closely related processes of the land surface energy balance as shown on the right hand side of Fig. 5-1.

5.1 Precipitation

As we saw in the previous chapter, one of the earliest cornerstones of soil physics was laid in 1856 by the French engineer, Henry Darcy, whose work focused on water filters and groundwater aquifers. Shortly thereafter a German, Ewald Wollny, became perhaps the first scientist whose published work focused squarely on soil physics. Wollny's experiments as early as 1874, were the first to quantitatively describe how raindrops falling on bare soil degraded the soil structure and how the vegetative canopy provided a vital protection against raindrop impact [1]. His subsequent experiments revealed that the vegetative canopies of common crops could intercept from 12-55% of the total rainfall. These discoveries, and his subsequent studies on runoff and erosion, make Wollny one of the pioneers of soil physics and hydrology and an early leader on the important issue of soil and water conservation. Our study of soil water balance process will begin where Wollny, and arguably soil physics itself, began – by considering precipitation and the interactions between raindrops, vegetative canopies, and the soil.

5.1.1 Precipitation Amount
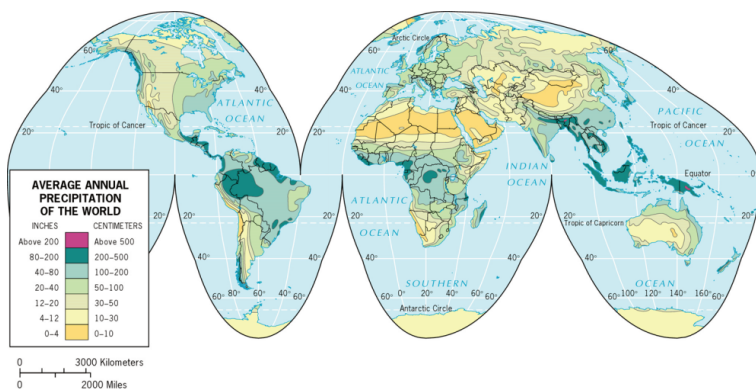
Fig. 5-2. Average annual precipitation across the Earth's land area. Source: planetolog.com (link).

The global pattern of precipitation amounts over land is arguably one of the most important influences on the Earth's coupled human and natural systems. Average annual precipitation ranges from <100 mm in the great deserts of the world to >2000 mm in the equatorial rainforests (Fig. 5-2). In fact, average annual rainfall exceeds 5000 mm (16 feet!) in a few locations, such as eastern India, where mountains lift warm, moist ocean winds cooling the air and resulting in precipitation. This process of orographic lift produces relatively high precipitation amounts on the upwind side of the mountains and relatively low precipitation areas, known as rain shadows, on the downwind side. The Earth's precipitation patterns are clearly reflected, not only in the global distribution of natural ecosystems, but in the patterns of human population, as well (Fig. 5-3). The majority of the world's cities with >1,000,000 people are located in regions with average annual rainfall between 500 and 2000 mm. If climate change produces substantial changes in precipitation patterns, then shifts in human population patterns or

increased water transfers by humans or both are likely to result. In fact, large scale water transfer projects have already begun, such as the enormous "South-to-North Water Diversion" project in China, which allows transfer of 25 x 109 (109 = billion) cubic meters of water per year from the Yangtze River basin in south China to Beijing and other locations in north China across a distance of more than 1,000 km [2].
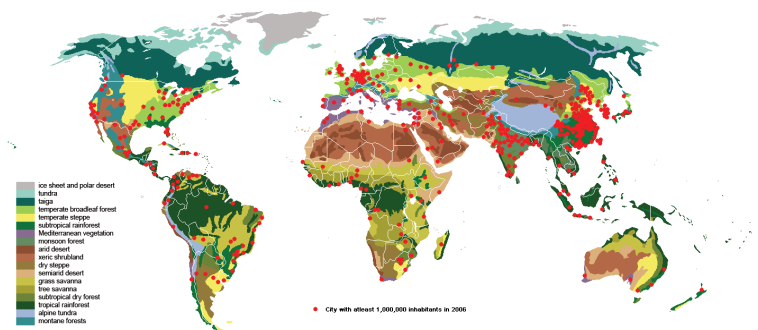
Fig. 5-3. Map of the Earth's terrestrial biomes and the locations of all cities with >1,000,000 people as of 2006. Author: KVDP. License: CCA–SA 3.0 (link).

5.1.2 Precipitation Intensity

The afternoon of June 22, 1947, was warm and humid in Holt, Missouri, a small farming community northeast of Kansas City, and a rainstorm, which had dropped 10-60 mm of rain across eastern Nebraska and Kansas, was approaching from the northwest [3]. The residents of Holt had no way of knowing that this storm would rapidly intensify and drop 305 mm (12 inches) of rainfall on their community in a span of only 42 minutes, setting a world record for precipitation intensity and filling their homes with up to 60 cm of water and mud [4]. This record rainfall intensity for a duration of 42 minutes is equivalent to 43.6 cm h^{-1}. This extremely high rainfall rate

is far greater than the rates at which water can enter, i.e. infiltrate, most soils.

The world record precipitation intensities for other durations are equally incredible. As much as 38 mm have fallen in a single minute (Table 5-1)! Clearly atmospheric processes are capable of generating precipitation intensities far greater than the capacity of the soil to transmit that water. And, the kinetic energy delivered to the land surface by these intense rainfall events can be substantial. For rainfall intensities of 10 cm h^{-1} the kinetic energy of the falling rain can approach 10,000 J m^{-2} h^{-1} [5]. Thus, Wollny was correct in recognizing the critical role of the vegetative canopy and plant residues in protecting the soil from the erosive energy of rainfall.

Table 5-1. World-record precipitation depths for durations ranging from 1 minute to 1 year. These data were current as of 2003 and were obtained from the US National Weather Service and reported by Galmarini et al. (2004).

Duration	Amount	Location	Date
	mm		
1 minute	38	Barot, Guadeloupe	26 Nov 1970
1 hour	401	Shangdi, Inner Mongolia, China	3 Jul 1975
1 day	1,825	Foc Foc, La Réunion	7-8 Jan 1966
1 month	9,300	Cherrapunji, India	Jul 1861
1 year	26,461	Cherrapunji, India	Aug 1860–Jul 1861

The record rainfall totals for durations greater than a few minutes far exceed the amount of water that the atmosphere can hold in one location, proving that extreme rainfall events depend on strong horizontal transport of water vapor into the rainstorm from the surrounding atmosphere [6]. If all the atmospheric water vapor above a given land area were condensed and deposited on the land surface so it could be measured, that depth of water would be equal to what atmospheric scientists call the ***total column water vapor***

or **total precipitable water**. The annual average of this precipitable water ranges from near 60 mm over portions of the equatorial oceans and the Amazonian rain forest to near 0 mm in the rain shadows of the Himalayas and Andes mountains (Fig. 5-4). Understanding that the atmosphere typically holds less than 60 mm of precipitable water at any one place and time makes the record precipitation intensities in Table 5-1 all the more incredible.
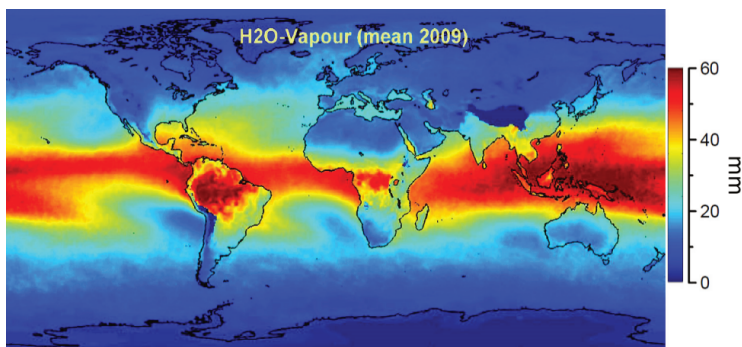
Fig. 5-4. Average total column water vapor in the atmosphere for 2009. Created by the European Space Agency Data User Element program GlobVapour project (link).

5.1.3   Raindrop Size Distribution

Closely related to rainfall intensity is raindrop size. Raindrops typically have cross-sectional diameters between 1 and 4 mm [7], and the size distribution shifts toward larger drops as the rainfall intensity increases (Fig. 5-5). The minimum diameter of raindrops is approximately 0.5 mm because smaller droplets are generally kept aloft by air currents. Although raindrops with equivalent diameters up to 9 mm have been observed [8], such large drops are generally unstable. The drag forces and pressures acting on a large drop as it falls result in the flattening of the drop, followed by deformation

into a short-lived, parachute-like shape, and ultimately the complete disintegration of the drop. This disintegration has been captured in striking high-speed video footage available here. In fact, this process of fragmentation of large drops may generate the entire distribution of raindrop sizes [9].
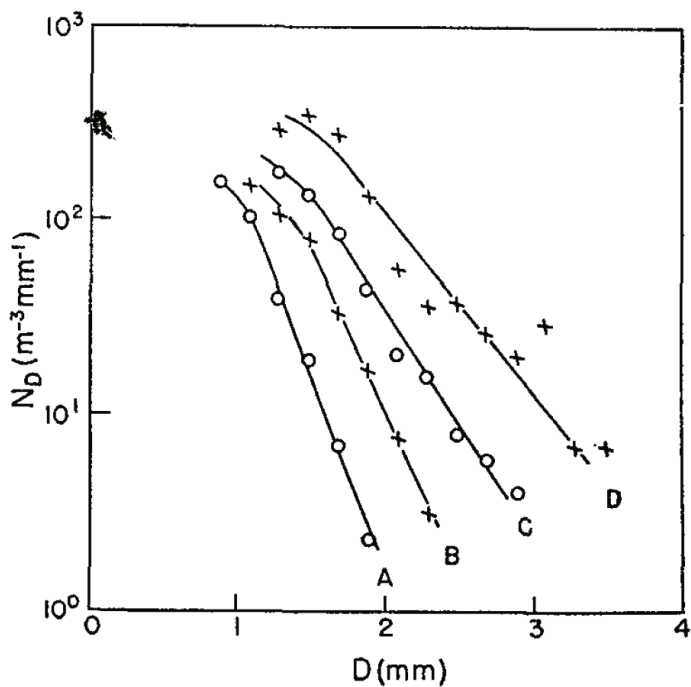
Fig. 5-5. Raindrop size distributions recorded at Ottawa, Canada. The curves A, B, C, and D represent rainfall intensities of 1.0, 2.8, 6.3, and 23.0 mm h^{-1}, respectively. The x-axis shows the drop diameter (D). The number of drops (N$_D$) on the y-axis is expressed such that N$_D$dD is the number of drops with diameters between D and D + dD in a unit volume of space. Reproduced from Marshall and Palmer (1948).

5.2 Rainfall Interception

In many terrestrial environments, before water inputs from rainfall or irrigation ever reach the soil surface, the vegetative canopy and plant residue alter both the amount and spatial distribution of those inputs. When rainfall is prevented from reaching the soil because of interception by plant canopies or plant residue and subsequent evaporation, that process is called *rainfall interception*. Related process include: throughfall, when rain passes through the canopy before reaching the soil; stemflow, when rain striking the plant canopy flows down leaves and branches to the main stem of the plant and flows down the stem to the soil; and canopy drip, when water from rain or dew drips off the outer edges of the plant canopy. Similar interception processes can occur when water inputs are supplied by sprinkler-type irrigation systems. These interception processes can strongly affect the soil water balance in both natural and managed ecosystems, and the magnitude of the effects are illustrated by the following case study.

5.2.1 Vegetation Characteristics Affecting Rainfall Interception

Soil water availability is typically the limiting factor for vegetative productivity of the expansive rangelands on the Edwards Plateau in central Texas, USA. Dominant species of vegetation in these rangelands include: sideoats grama [Bouteloua curtipendula (Michx.) Torr.], a bunchgrass; curlymesquite [Hilaria belangeri (Steud.) Nash] a shortgrass; and live oak (Quercus virginiana Mill.), a broadleaf, evergreen tree [10]. These species have different growth habits as illustrated in Fig. 5-6, and as a result, rainfall interception differs between the species.

Using a combination of field and laboratory measurements, researchers discovered that curlymesquite canopies intercepted more rainfall than sideoats grama in this environment on a *per unit mass basis* because curlymesquite has flat leaf blades covered with fine hair and also stolons, which grow horizontally at the soil surface [10]. These characteristics made curlymesquite relatively effective at intercepting rainfall. However, sideoats gramma intercepted more rainfall than curlymesquite on a *per unit area basis* because the mass of standing live and dead biomass was greater in sites dominated by sideoats grama than in sites dominated by curlymesquite.
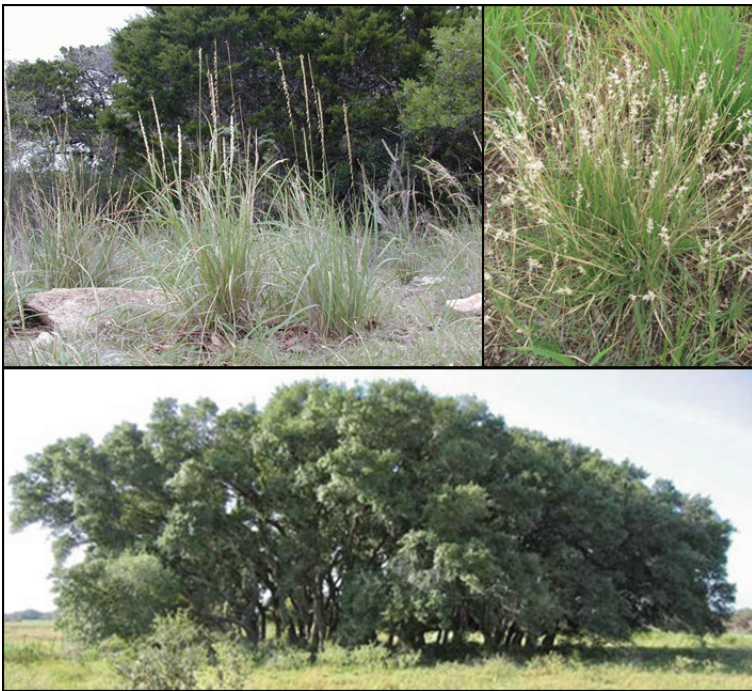
Fig. 5-6. Sideoats grama (top left; Source: Texas A&M AgriLife, link), curlymesquite (top right; Source: San Antonio River Authority, link), and live oak (bottom; Source: Texas A&M Forest Service, link).

The maximum volume of rainfall per unit area that a particular plant canopy can hold is the ***interception storage capacity*** of the canopy. For the sideoats grama-dominated site the interception storage capacity was 1.8 mm compared to 1.0 mm for the curlymesquite-dominated site. Based on the observed size distribution of rainfall events in the region, 18.1% of annual rainfall would be intercepted at the sideoats grama-dominated sites, compared with 10.8% at the curlymesquite dominated sites.

These interception totals are certainly enough to reduce water availability, but they were far exceeded by the rainfall interception of the clumps, or mottes, of live oak. In the live oak mottes, the canopy was dense, and the soil surface was covered by a thick accumulation of leaf litter. The live oak leaf litter could hold approximately twice as much rainfall per unit mass as the grasses and had an estimated interception storage capacity of 8.7 mm. Thus, the canopy and the leaf litter together were highly effective in rainfall interception, and only 53.9% of the annual precipitation reached the soil (Table 5-2).

Table 5-2. Partitioning of annual rainfall within live oak mottes based on 10-year average rainfall event size distribution for a site on the Edwards Plateau in central Texas, USA. Adapted from Thurow et al. (1987).

	Amount	Percent of Annual Rainfall
	mm	%
Annual Precipitation	523	
Canopy Interception	133	25.4
Throughfall + Stemflow	390	74.6
Litter Interception	108	20.7
Water Reaching the Soil	282	53.9

Although stemflow accounted for only 3.3% of annual precipitation, stemflow delivered a volume of water equivalent to 222% of the annual precipitation to the soil within a 10-cm radius of each tree. This redistribution of water by the trees may help them establish a competitive advantage over other plant species, and similar competitive mechanisms have been suggested for juniper tree encroachment in rangeland environments [11, 12].

Rainfall interception can substantially reduce soil water availability not only in rangeland environments but also in forests and croplands. Interception has been reported to account for 33% of rainfall in winter wheat (*Triticum aestivum* L.), 30% in maize (*Zea mays* L.), 35% in soybean [*Glycine max* (L.) Merr.], 25-31% in switchgrass (*Panicum virgatum* L.), and 27-45% in forage sorghum (*Sorghum bicolor* L. Moench) [13-15]. Interception also occurs during sprinkler irrigation and accounted for approximately 8% of applied water for a maize crop irrigated with a center pivot system in western Kansas, USA [16]. Other examples of rainfall interception percentages in natural ecosystems include 36% under eastern redcedar (*Juniperus virginiana* L.), 44% under tallgrass prairie, 9% under Amazonian rainforest, and 20-29% in a deciduous forest [17-19].

5.2.2 Rainfall Characteristics Affecting Interception

Rainfall interception is influenced, not only by the vegetation and residue characteristics, but also by the properties of the rainfall itself. Rainfall events can be characterized by their amount, intensity, and duration, with the amount of rainfall per event evidently exerting the strongest influence of the three on rainfall interception [10]. In the case of interception in live oak mottes on the Edwards Plateau, the data show that small precipitation events <5 mm were completely intercepted by the canopy and the leaf litter (Fig. 5-7) with no rainfall reaching the soil. Events totals >15 mm were required to ensure that at least half of the precipitation reached the soil, and even for the largest events, ~80 mm, only about 80% of the rainfall reached the soil. Therefore, regions which receive a relatively large portion of their precipitation through small rainfall events will be susceptible to relatively large losses to due rainfall interception.
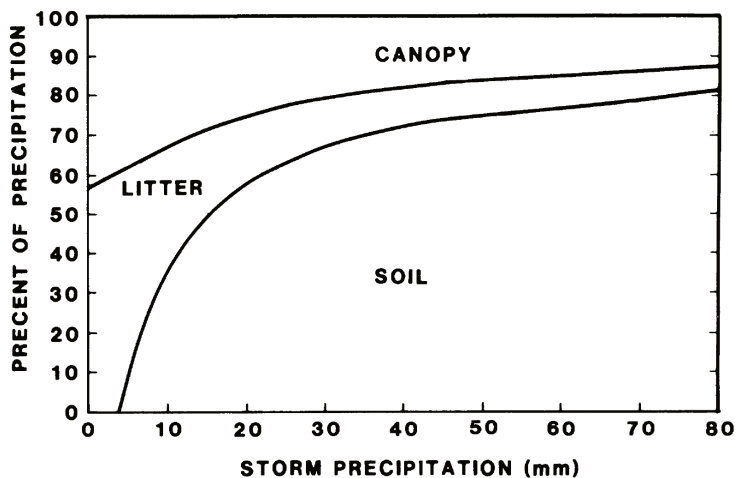
Fig. 5-7. Percent of precipitation intercepted by the plant canopy or leaf litter or reaching the mineral soil under live oak mottes on the Edwards Plateau in central Texas, USA, as a function of the total amount of precipitation in an individual event or storm. Reproduced from Thurow et al. (1987).

Rainfall interception may also be influenced by the raindrop size distribution [20]. Small raindrops are more effective than large drops in wetting the surfaces of vegetation or residue, and small drops have less kinetic energy than large drops, which increases the likelihood of small drops being retained by the vegetation or residue. As a result, interception losses tend to increase as the raindrop size decreases.

5.3 Raindrop Impact

Raindrops whose fall is not interrupted by vegetation or plant residues can strike the soil surface like miniscule bombs going off. The greater the speed and mass of the drop, the greater the force of the impact. The maximum speed a given drop can reach during its fall, its *terminal velocity*, increases as the drop size increases. Small water droplets in the atmosphere with diameters of 0.1 mm have terminal velocities of only 0.27 m s^{-1}, while large raindrops with diameters of 5.8 mm have terminal velocities of 9.2 m s^{-1} [21]. When raindrops strike the soil surface at high speed, soil can be splashed into the air by these drop impacts, and this *soil splash* contributes significantly to the spread of plant pathogens and to soil erosion and soil crust formation, processes which we will consider in subsequent sections [22, 23].

Fig. 5-8. Photograph of soil splash from a single raindrop. Image credit: US Department of Agriculture (link).

The amount of soil splash generated by raindrop impact is influenced, not only by the drop size and speed, but also by the strength of the soil. **Soil strength** is the maximum stress that a particular soil body can bear without failing. When a raindrop impacts high strength soil, the drop does not appreciably penetrate the soil and the drop spreads radially across the soil surface, resulting in relatively little soil splash [24]. In contrast, when a drop first impacts low strength soil, the soil surface is deformed under the drop, raising a ring-shaped bulge of soil around the edge of the drop (Fig. 5-9). Assuming that the soil is saturated, the soil volume and density are unchanged because there is insufficient time during the impact for the soil to drain and water is essentially incompressible. As the impact continues, the force of the spreading water blows out the sides of the cavity splashing soil into the air. Lower soil strength results in a larger cavity and surrounding bulge, greater soil detachment, and greater splash angle relative to the soil surface. This soil detachment often initiates the processes of soil crusting, which we will consider here, and soil erosion, which we will consider in a later chapter.
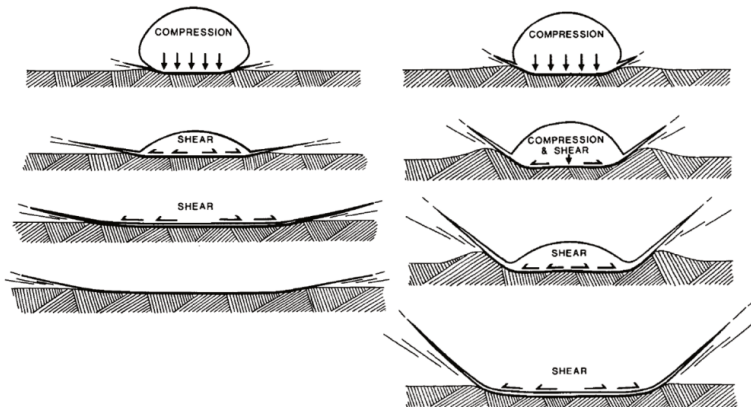
Fig. 5-9. Conceptual diagram of soil splash process for high strength (left) and low strength (right) soil. Adapted from Al-Durrah and Bradford (1982).

5.4 Crust Formation
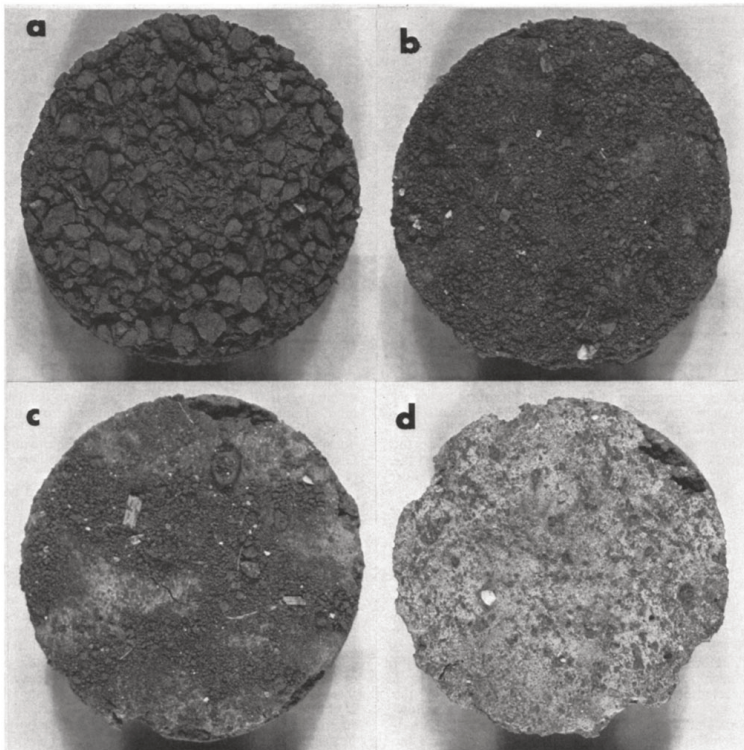
Fig. 5-10. Photographs of a silty clay loam (Capay soil series) subjected to simulated rainfall at 42.5 mm h^{-1} with a 3.2 mm drop diameter and 5.9 m s^{-1} drop velocity. Photos were taken after 0 (a), 10 (b), 25 (c), and 90 (d) minutes of rainfall. Samples are 5-cm diameter. Reproduced from Moore and Singer (1990).

A **soil crust** is a relatively thin layer at the soil surface which has a higher bulk density, a reduced number of large pores, and a lower hydraulic conductivity than the underlying soil. Fig. 5-10 illustrates

the dramatic changes that can occur at the soil surface during raindrop impact and crust formation [25]. During the first stage of crust formation under simulated rainfall (panel b), the aggregates at the surface of this silty clay loam soil were broken down into smaller particles as their soil water content increased and soil strength declined. In the second stage of crust formation (panel c), much of the relatively fine material resulting from the aggregate breakdown was removed from the surface by soil splash and runoff. Some soil particles also flowed with the infiltrating water into the pores of the surface layer, where they were deposited thus lowering the porosity. In the third and final stage of crust formation (panel d), the surface was devoid of loose material, a cohesive surface layer had formed, and the rate of soil splash declined.
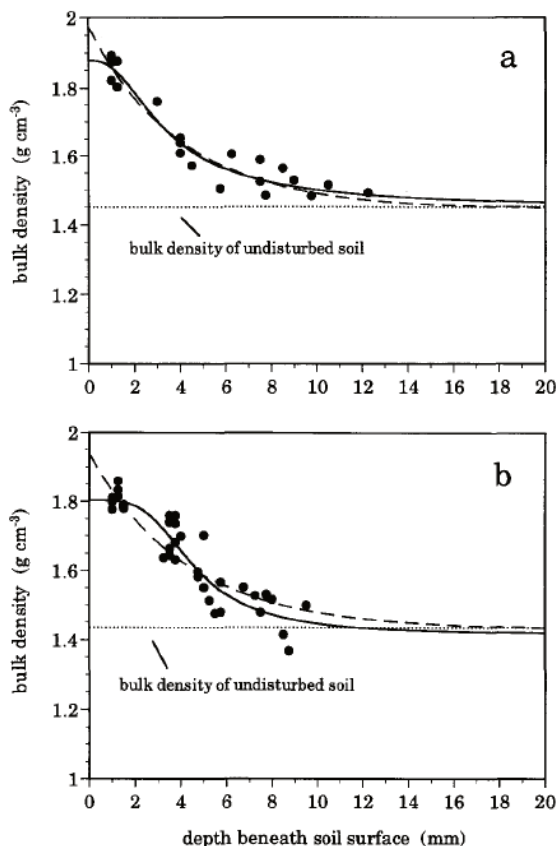
Fig. 5-11. Bulk density versus depth for a loamy sand (a) and sandy loam (b) subjected to 60 mm of simulated rain. Adapted from Roth (1997).

These crust forming processes alter, not only the soil surface, but a layer which can extend >1 cm beneath the surface. For example, 2 h of simulated rainfall at a rate of 30 mm h^{-1} resulted in a zone of increased bulk density extending to a depth of ~15 mm for a loamy sand and a sandy loam from Germany [26]. The bulk density reached maximum values of approximately 1.90 g cm^{-3} near the soil surface, an exceptionally high density relative to the 1.45 g cm^{-3} of

the underlying soil (Fig. 5-11). The high bulk density, low porosity, and the absence of large pores in the crust can greatly reduce the hydraulic conductivity of the surface soil. For instance, the hydraulic conductivity of the surface soil was reduced to <20% of its initial value following 20 minutes of simulated rain at 50 mm h^{-1} on four soils from Georgia that were susceptible to crusting [27].

Soil crusts have a diverse array impacts. Seedling emergence can be delayed or prevented entirely by soil crusts (Fig. 5-12). One way to characterize a soil crust is by its **penetration resistance**, which is an indicator of soil strength and can be measured relatively easily with a hand-held penetrometer (link). As the penetration resistance of soil crusts increased from 300 kPa to 900 kPa, cotton seedlings' emergence was reduced from ~60% to ~25% in one study in Texas [28]. Soil crusts can also reduce the rate of infiltration into the soil.

Fig. 5-12. Seedling breaking through a soil crust. Source: Iowa State University Extension (link).

Two primary mechanisms of soil crust formation are physical

dispersion and chemical dispersion of soil aggregates. Chemical dispersion and the factors contributing to chemical dispersion have been discussed in section 4.5. ***Physical dispersion*** is the breakdown of soil aggregates caused by raindrop impact, by rapid rewetting of initially dry aggregates (recall Fig. 2-9), or by mechanical disturbance such as tillage. The degree of physical dispersion is influenced by both the severity of the disruptive forces and by the stability of the soil aggregates.

5.5 Problem Set

1. Calculate the average volumetric water content after rainfall event for a 68 cm deep soil profile which had an average volumetric water content of 0.27 cm^3 cm^{-3} before getting 6.4 cm of rain. Assume 28% of the rain was lost to interception and there were no other losses.

5.6 References

1. Baver, L., *Ewald Wollny—A pioneer in soil and water conservation research.* Soil Science Society of America Journal, 1938. **3**: p. 330-333.

2. Barnett, J., et al., *Sustainability: transfer project cannot meet China's water needs.* Nature News, 2015. **527**(7578): p. 295.

3. Lott, G.A., *The world-record 42-minute Holt, Missouri, rainstorm.* Mon. Wea. Rev, 1954. **82**(2): p. 50-59.

4. Locatelli, J.D. and P.V. Hobbs, *A world record rainfall rate at Holt, Missouri: Was it due to cold frontogenesis aloft?* Weather and forecasting, 1995. **10**(4): p. 779-785.

5. Salles, C., J. Poesen, and D. Sempere-Torres, *Kinetic energy of rain and its functional relationship with intensity.* Journal of Hydrology, 2002. **257**(1–4): p. 256-270.

6. Galmarini, S., D.G. Steyn, and B. Ainslie, *The scaling law relating world point-precipitation records to duration.* International Journal of Climatology, 2004. **24**(5): p. 533-546.

7. Marshall, J.S. and W.M.K. Palmer, *The distribution of raindrops with size.* Journal of meteorology, 1948. **5**(4): p. 165-166.

8. Fujiyoshi, Y., et al. *The maximum size of raindrops—can it be a proxy of precipitation climatology.* Int. Conf. Clouds and Precipitation, Cancun, Mexico. 2008.

9. Villermaux, E. and B. Bossa, *Single-drop fragmentation determines size distribution of raindrops.* Nature Physics, 2009. **5**(9): p. 697-702.

10. Thurow, T.L., et al., *Rainfall Interception by Midgrass, Shortgrass, and Live Oak Mottes.* Journal of Range Management, 1987. **40**(5): p. 455-460.

11. Robinson, D.A., et al., *Soil Water Repellency: A Method of Soil Moisture Sequestration in Pinyon-Juniper Woodland.* Soil Science Society of America Journal, 2010. **74**(2): p. 624-634.

12. Wine, M.L., et al., *Effects of eastern redcedar encroachment on*

soil hydraulic properties along Oklahoma's grassland-forest ecotone. Hydrological Processes, 2012. **26**(11): p. 1720-1728.

13. Haynes, J., *Ground rainfall under vegetative canopy of crops.* Journal of the American Society of Agronomy, 1940. **32**: p. 176-184.

14. Leuning, R., et al., *Rainfall interception and evaporation from soil below a wheat canopy.* Agricultural and Forest Meteorology, 1994. **67**(3-4): p. 221-238.

15. Yimam, Y.T., T.E. Ochsner, and V.G. Kakani, *Evapotranspiration partitioning and water use efficiency of switchgrass and biomass sorghum managed for biofuel.* Agricultural Water Management, 2015. **155**(0): p. 40-47.

16. Steiner, J., E. Kanemasu, and R. Clark, *Spray losses and partitioning of water under a center pivot sprinkler system.* Transactions of the ASAE, 1983. **26**(4): p. 1128-1134.

17. Herbst, M., et al., *Seasonal variability of interception evaporation from the canopy of a mixed deciduous forest.* Agricultural and Forest Meteorology, 2008. **148**(11): p. 1655-1667.

18. Lloyd, C.R., et al., *The measurement and modelling of rainfall interception by Amazonian rain forest.* Agricultural and Forest Meteorology, 1988. **43**(3): p. 277-294.

19. Zou, C.B., et al., *Canopy Interception for a Tallgrass Prairie under Juniper Encroachment.* PLOS ONE, 2015. **10**(11): p. e0141422.

20. Calder, I.R., *Dependence of rainfall interception on drop size: 1. Development of the two-layer stochastic model.* Journal of Hydrology, 1996. **185**(1): p. 363-378.

21. Gunn, R. and G.D. Kinzer, *The terminal velocity of fall for water droplets in stagnant air.* Journal of Meteorology, 1949. **6**(4): p. 243-248.

22. Salles, C. and J. Poesen, *Rain properties controlling soil splash detachment.* Hydrological Processes, 2000. **14**(2): p. 271-282.

23. Madden, L., *Effects of rain on splash dispersal of fungal pathogens.* Canadian Journal of Plant Pathology, 1997. **19**(2): p. 225-230.

24. Al-Durrah, M. and J. Bradford, *The mechanism of raindrop splash on soil surfaces.* Soil Science Society of America Journal, 1982. **46**(5): p. 1086-1090.
25. Moore, D.C. and M.J. Singer, *Crust Formation Effects on Soil Erosion Processes.* Soil Science Society of America Journal, 1990. **54**(4): p. 1117-1123.
26. Roth, C.H., *Bulk density of surface crusts: depth functions and relationships to texture.* Catena, 1997. **29**(3): p. 223-237.
27. Chiang, S.C., D.E. Radcliffe, and W.P. Miller, *Hydraulic Properties of Surface Seals in Georgia Soils.* Soil Science Society of America Journal, 1993. **57**(6): p. 1418-1426.
28. Bilbro, J. and D. Wanjura, *Soil crusts and cotton emergence relationships.* Transactions of the ASAE, 1982. **25**(6): p. 1484-1487.

6. INFILTRATION

In the prior chapter, we considered water inputs to the soil and the ways in which those water inputs can be altered due to interception. We also looked at the process of raindrop impact on the soil surface and its potentially harmful effects. Now, we are prepared to go-with-the-flow and consider the next step in the soil water balance–the process of infiltration. **Infiltration** is simply the process by which water enters the soil profile. Infiltration was the focus of perhaps one of the first and certainly one of the most influential soil physics teaching videos ever produced. That classic video produced at Washington State University in 1959 by Walter H. Gardner and his associate, J.C. Hsieh, inspired a modern remake created at the University of Gembloux in Belgium and available here (link). Please take the time now to watch that video, which will help you better visualize and understand key features of the infiltration process.

The **infiltration rate** is the volume of water flowing into the soil per unit of surface area per unit time. As seen in the video above, infiltration rates typically decrease over time during the course of an infiltration event, approaching a relatively constant and low rate if the infiltration event is prolonged. You can see in Fig. 6-1 how the infiltration rates initially declined sharply then approached a fairly constant level during ponded infiltration into a silty clay loam soil in Hawaii [1]. The fundamental cause of this decrease in infiltration rates over time was one of the main scientific mysteries in the early days of the soil physics and hydrology. This mystery was studied in detail by an American engineer named Robert Horton, one of the most prominent early hydrologists.
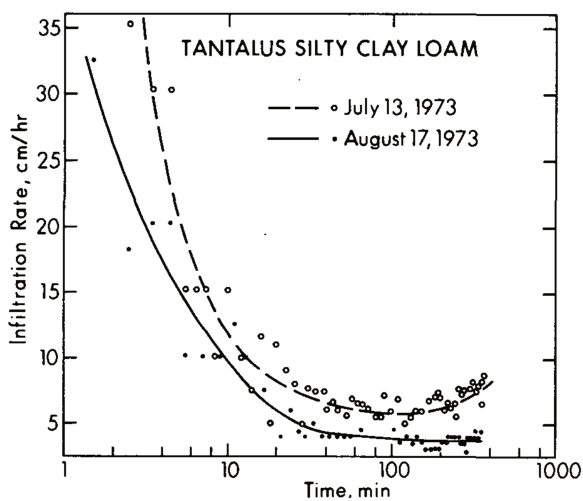
Fig. 6-1. Infiltration rates for a Tantalus silty clay loam in Hawaii measured using a double-ring infiltrometer. Adapted from Ahuja et al. (1976).

6.1 Horton Infiltration Model

Horton defined infiltration capacity as "the maximum rate at which a given soil can absorb rainfall when the soil is in a specified condition" [2]. Based on observations of infiltration in the field, he deduced that the decrease in infiltration capacity over time during an infiltration event was due to flow-restricting changes occurring in a thin layer at the soil surface. He believed the key changes included: 1) packing of the soil surface by raindrop impact, 2) swelling of the soil, and 3) plugging of surface pores with fine materials [2]. Building on this conceptual framework, he developed what came to be known as the Horton infiltration equation:

$$f = f_c + 9f_o + f_c)_e^{-K_f t}$$

(Eq. 6-1)

where f is the infiltration capacity, f_c is a constant minimum value of infiltration capacity which f approaches asymptotically, f_0 is the soil's initial infiltration capacity at the start of the infiltration event, and K_f is a constant which controls the rate at which the infiltration capacity decreases as a function of time, t [3]. This model has been widely used in hydrology and related disciplines [4], and its success may be due to the correctness of its underlying conceptual framework or the flexibility of its mathematical form or some combination of the two. The main deficiencies of this infiltration model are that the parameters are not clearly related to measurable physical properties of the soil and, more importantly, that the model's conceptual framework fails to include the most fundamental reason that infiltration rates decrease over time. That fundamental discovery had, in fact, been made decades before but apparently was not widely accepted in Horton's time.
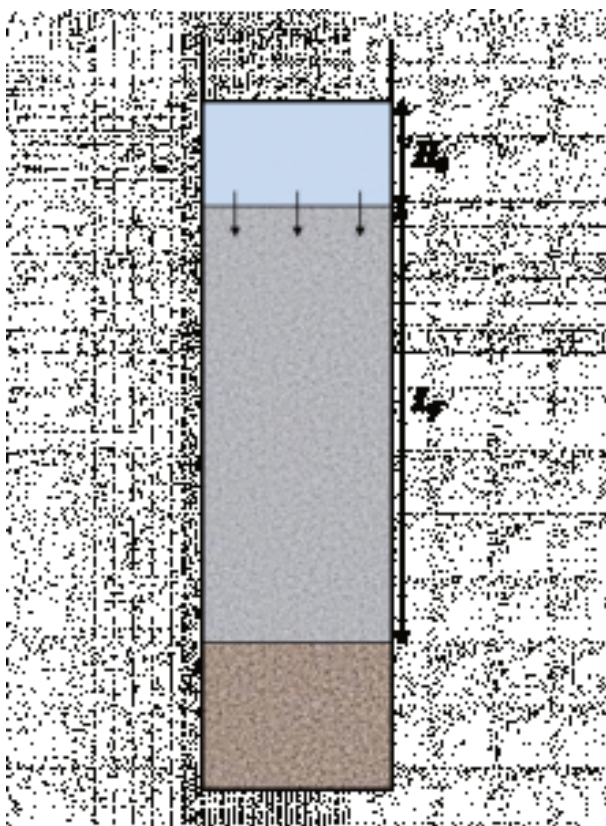
Fig. 6-2. Uniform soil column undergoing downward infiltration with constant ponding of depth H_0 on the surface and a distinct wetting front located at a depth L_f below the surface.

6.2 Green-Ampt Infiltration Model

In the early 1900s, Heber Green and G.A. Ampt were studying drainage and soil water flow in Australia. They built their efforts on the early soil physics literature, including the seminal contribution of Edgar Buckingham on flow in unsaturated soil, as described in section 4.7. In 1911, they published physically-based equations describing downward infiltration, upward infiltration (as from a water table), and horizontal infiltration (as from a channel or ditch) [5]. For the case of downward infiltration, the flow situation envisioned by Green and Ampt is represented in Fig. 6-2. Their approach resulted in a simplistic but elegant approximation for infiltration when water is ponded on the soil surface.

The Green-Ampt model for downward infiltration can be written as:

$$i = K \frac{(L_f + H_0 - H_f)}{L_f}$$

<div align="right">(Eq. 6.2)</div>

where i is the infiltration rate, K is the hydraulic conductivity of the soil, L_f is the distance (or length) from the soil surface to the wetting front, H_0 is the pressure head at the soil surface, and H_f is the pressure head at the wetting front. This model is based on the following assumptions:

- homogeneous soil with a uniform initial water content, θ_i
- the pressure head at the wetting front, H_f, is constant
- the water content and hydraulic conductivity, K, are uniform and constant in the wetted region
- the ponding depth, H_0, is constant

This model works best when a relatively sharp or distinct wetting front exists throughout the infiltration process. Such a distinct wetting front is more likely to occur in coarse-textured soils than in fine-textured soils and in initially dry soil than in initially wet soil. The model also works best when the soil texture is homogeneous throughout the wetted region and when air-entrapment, surface crusting, and soil swelling do not substantially influence the infiltration process.

The Green-Ampt model also provides a way to estimate cumulative infiltration during an infiltration event. The relevant equation is

$$I = L_f \cdot \Delta\Theta$$

(Eq. 6-3)

where I is the cumulative infiltration and $\Delta\theta$ is the difference between the final water content, θ_f, and the initial water content, θ_i. One difficulty in applying this equation is that there is not a direct (i.e. explicit) way to calculate how the position of the wetting front, L_f, changes over time, t. For that purpose, we have to solve the following equation for L_f by trial-and-error (i.e. implicitly):

$$\frac{K \cdot t}{\Delta\Theta} = L_f - (H_o - H_f) \ln\left(1 + \frac{L_f}{H_o - H_f}\right)$$

(Eq. 6-4)

Despite its assumptions and caveats, the Green-Ampt model has proven to be extremely useful for helping us understand and predict infiltration. The power of this model lies in the fact that it succinctly describes the most fundamental reason why infiltration rates tend to decrease over time–*the hydraulic gradient driving infiltration tends to decrease over time.* The Green-Ampt model can be viewed as a specialized application of the Buckingham-Darcy law (Eq. 4-5) as explained in this video (link). The ratio on the right hand side of the Green-Ampt equation represents the hydraulic gradient. Since the ponding depth, H_0, and the pressure potential at the wetting front, H_f, are treated as constants, the hydraulic gradient gets

smaller as L_f gets larger, i.e. as the wetting front moves farther away from the soil surface.
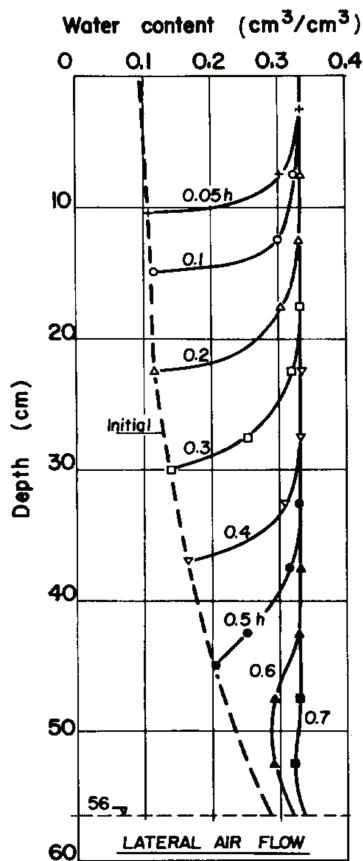
Fig. 6-3. Water content profiles during ponded infiltration (H0 = 1.5 cm) into a fine sand. Adapted from Vachaud et al. (1974).

The main features of the infiltration process as represented by the Green-Ampt model are evident in experimental results, such as

those shown in Fig. 6-3. These data from a laboratory experiment using a soil column filled with fine sand clearly show how the rate of advance of the wetting front, and thus the infiltration rate, decreases over time [6]. The wetting front reached the 10 cm depth in only about 0.05 h (3 minutes), but it took about four times as long (0.2 h or 12 minutes) to reach the 20 cm depth. Thus, these data also emphasize a fundamental tendency of infiltration−*during the early part of an infiltration event, the wetting front position is approximately proportional to the square root of the event duration.* Cumulative infiltration follows this same proportionality, and the Green-Ampt model captures this behavior.

6.3 Infiltration for a Constant Rainfall Rate

The Green-Ampt model described in the prior section is only valid for ponded infiltration. It does not provide infiltration predictions for the case when water is supplied to the soil surface, by rain or sprinkler irrigation, at a rate below the soil's infiltration capacity. In 1971, engineers Russell Mein and Curtis Larson at the University of Minnesota made an important breakthrough by modifying the Green-Ampt model to account for the case of infiltration under constant-rate rainfall [7]. They reasoned that if rain is reaching the soil surface at a constant rate that is below the soil's initial infiltration capacity but above the soil's saturated hydraulic conductivity then initially all the water will infiltrate as it reaches the surface. During that stage, infiltration is **_supply controlled_**, i.e. the rate of water delivery to the surface limits the infiltration rate. However, since the soil's infiltration capacity decreases over time during an event, a distinct point in time will be reached when the infiltration capacity falls below the rainfall rate and water will begin to accumulate, or pond, on the surface. After this time, infiltration is **_soil controlled_**, i.e. the rate of water flow in the soil limits the infiltration rate. Only after this change point is reached can runoff begin. Study Fig. 6-4 for a graphical representation of this conceptual model for the infiltration process.
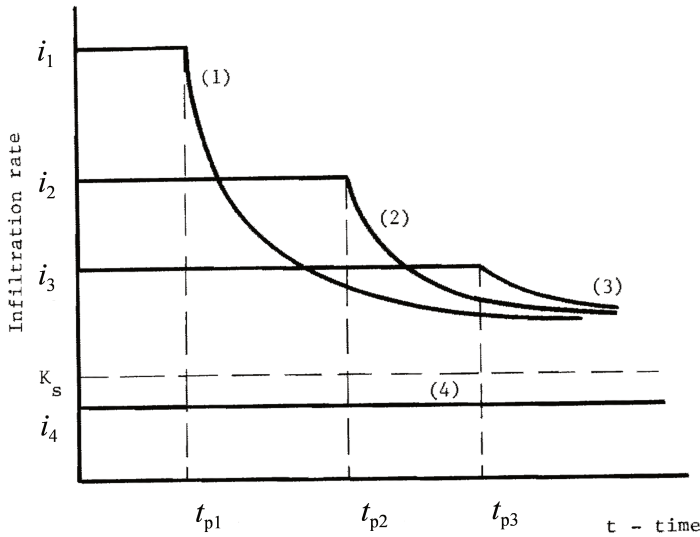
Fig. 6-4. Hypothetical curves of infiltration rate, i, versus time, t, during infiltration of rainfall reaching the soil surface at a constant rate. The solid curves labeled (1) through (4) represent different rainfall rates, with curve 1 being the greatest and curve 4 the least. Curve 4 represents a rainfall rate below the soil's saturated hydraulic conductivity, K_s. The end of the horizontal segment of each curve indicates the time at which water first begins to pond on the surface, t_p. Adapted from Mein and Larson (1971).

The Mein and Larson modification of the Green-Ampt model provides a simple three-step way to estimate this time to ponding. The first step is to note that ponding begins the moment when the rainfall rate, r, equals the soil's infiltration capacity, i, as defined by the Green-Ampt ponded infiltration equation, Eq. 6-2. Substituting r for i in Eq. 6-2, assuming $H_0 = 0$, and solving for L_f gives:

$$L_f = \frac{-H_f}{\frac{r}{K} - 1}$$

(Eq. 6-5)

which defines the depth of the wetting front when ponding

begins. Next, the cumulative infiltration, I, that has occurred prior to ponding is estimated by inserting this L_f value in Eq. 6-3. Finally, the time required for the rainfall rate, r, to supply the cumulative water depth, I, is estimated by:

$$t_p = \frac{I}{r}$$

(Eq. 6-6)

where t_p is the time required for ponding to occur.

6.4 Infiltration Measurements

Fig. 6-5. Photographs of a double-ring infiltrometer (top left; Source: Wikipedia, public domain), single-ring infiltrometer (top right; Source: Wikipedia, public domain), mini-disk tension infiltrometer (bottom right), and sprinkle infiltrometer (bottom left, Source: Iowa Learning Farms, used by permission).

Infiltration can be measured under controlled conditions using devices called **infiltrometers**. Double-ring and single-ring infiltrometers are simple metal rings pushed or pounded into the soil surface (Fig. 6-5). Water is applied to the soil surface inside the rings either with a **constant head**, i.e. depth that is held constant or nearly constant by continual refilling of the ring as water infiltrates, or a **falling head**, i.e. the depth of ponded water is allowed to decrease over time as water infiltrates. With the single-ring

infiltrometer, the infiltrating water flows laterally and vertically away from the ring, and the effects of the lateral flow must be accounted for in order to meaningfully interpret the infiltration measurements and to estimate soil properties governing infiltration, such as the soil hydraulic conductivity. The outer ring of the double-ring infiltrometer is intended to reduce lateral flow from the inner ring, allowing the data from the inner ring to be analyzed as if the flow were one-dimensional. When proper field procedures and analytical techniques are applied, both the single- and double-ring infiltrometers provide similar estimates of hydraulic conductivity [8].

Unlike ring infiltrometers which apply water to the soil surface at positive pressures, tension infiltrometers like the one shown in Fig. 6-5 apply water under slightly negative pressures, often between -1 and -10 cm of water. Tension infiltrometers are portable, require relatively short measurement times, and require smaller volumes of water than do single- or double-ring infiltrometers. Tension infiltrometers may also result in less variable hydraulic conductivity estimates than other methods, but those estimates are, in some cases, lower than those from other methods [9].

Another type of infiltrometer, often called a sprinkle infiltrometer, applies water to the soil as drops from a collection of capillary tubes suspended a few centimeters above the surface [10]. The water delivery rate can be controlled by adjusting the pressure inside the sealed water supply reservoir and measured by the decrease of the water height in the reservoir (Fig. 6-5). The water application continues until runoff is generated beneath the infiltrometer. The runoff is collected and measured, and the infiltration rate is calculated as the difference between the water delivery rate and the runoff rate [11].

Fig. 6-6. Photograph of a 20-m long by 5- m wide runoff plot in Queensland, Australia. This plot is equipped with a tipping bucket device to measure runoff. Reproduced from Carroll et al. (2000).

Infiltration under natural rainfall conditions can be measured using runoff plots, which are small areas of land surrounded by borders to keep out surface runoff from adjacent areas and having equipment to collect and measure runoff on their downhill side. Infiltration can be estimated as the difference between precipitation and runoff. Fig. 6-6 shows a runoff plot used to measure infiltration and runoff at an open-cut coal mine site in Australia [12]. When the coal mine spoils lacked vegetation, they were susceptible to raindrop impact and crust formation, which resulted in low infiltration and high runoff volumes.

6.5 Problem Set

1. Water is ponded at a constant depth of 2.5 cm on a silt loam soil which had an initial water content of 0.15 cm^3 cm^{-3}. The water content of the wetted soil above the wetting front is 0.49 cm^3 cm^{-3}, and the wetting front has a pressure head of -17 cm. The hydraulic conductivity of the wetted soil is 0.65 cm h^{-1}. Apply the Green–Ampt model for ponded infiltration to answer the following questions:

   a. What is the infiltration rate (cm h^{-1}) when the wetting front reaches a depth of 30. cm?

   b. In total, how much water (cm) has infiltrated by that time?

   c. How long did it take (h) for the wetting front to reach that depth?

2. For that same soil with the same initial condition, consider the infiltration process for rainfall at a constant rate of 1.5 cm h^{-1}. Use the Mein and Larson modification of the Green–Ampt model to answer the following:

   a. How deep is the wetting front (cm) when ponding begins?

   b. In total, how much water (cm) has infiltrated by that time?

   c. How long did it take (h) before ponding occurred?

3. Repeat problem 2 but this time use a hydraulic conductivity of 0.325 cm h^{-1}, exactly half of the value used previously.

   a. How deep is the wetting front (cm) when ponding begins?

   b. In total, how much water (cm) has infiltrated by that time?

   c. How long did it take (h) before ponding occurred?

6.6 References

1. Ahuja, L.R., S.A. El-Swaify, and A. Rahman, *Measuring Hydrologic Properties of Soil with a Double-ring Infiltrometer and Multiple-depth Tensiometers1.* Soil Science Society of America Journal, 1976. **40**(4): p. 494-499.

2. Horton, R.E., *The role of infiltration in the hydrologic cycle.* Eos, Transactions American Geophysical Union, 1933. **14**(1): p. 446-460.

3. Horton, R.E., *Analysis of runoff-plat experiments with varying infiltration-capacity.* Eos, Transactions American Geophysical Union, 1939. **20**(4): p. 693-711.

4. Chahinian, N., et al., *Comparison of infiltration models to simulate flood events at the field scale.* Journal of Hydrology, 2005. **306**(1–4): p. 191-214.

5. Green, W.H. and G. Ampt, *Studies on Soil Phyics.* The Journal of Agricultural Science, 1911. **4**(01): p. 1-24.

6. Vachaud, G., J.P. Gaudet, and V. Kuraz, *Air and water flow during ponded infiltration in a vertical bounded column of soil.* Journal of Hydrology, 1974. **22**(1–2): p. 89-108.

7. Mein, R.G. and C.L. Larson, *Modeling the infiltration component of the rainfall-runoff process*, in *Water Resource Research Center Bulletin 43.* 1971, University of Minnesota.

8. Verbist, K., et al., *Comparison of Single- and Double-Ring Infiltrometer Methods on Stony Soils.* Vadose Zone Journal, 2010. **9**(2): p. 462-475.

9. Reynolds, W.D., et al., *Comparison of Tension Infiltrometer, Pressure Infiltrometer, and Soil Core Estimates of Saturated Hydraulic Conductivity.* Soil Science Society of America Journal, 2000. **64**(2): p. 478-484.

10. Ogden, C.B., H.M. van Es, and R.R. Schindelbeck, *Miniature Rain Simulator for Field Measurement of Soil Infiltration.* Soil Science Society of America Journal, 1997. **61**(4): p. 1041-1043.

11. Alizadehtazi, B., et al., *Comparison of Observed Infiltration Rates of Different Permeable Urban Surfaces Using a Cornell Sprinkle Infiltrometer.* Journal of Hydrologic Engineering, 2016. **21**(7): p. 06016003.

12. Carroll, C., L. Merton, and P. Burger, *Impact of vegetative cover and slope on runoff, erosion, and water quality for field plots on a range of soil and spoil materials on central Queensland coal mines.* Soil Research, 2000. **38**(2): p. 313-328.

7. RUNOFF AND WATER EROSION

Having considered the process of infiltration in the prior chapter, we are now ready to move to the next processes in the soil water balance–the interconnected processes of runoff and water erosion (Fig. 7-1).
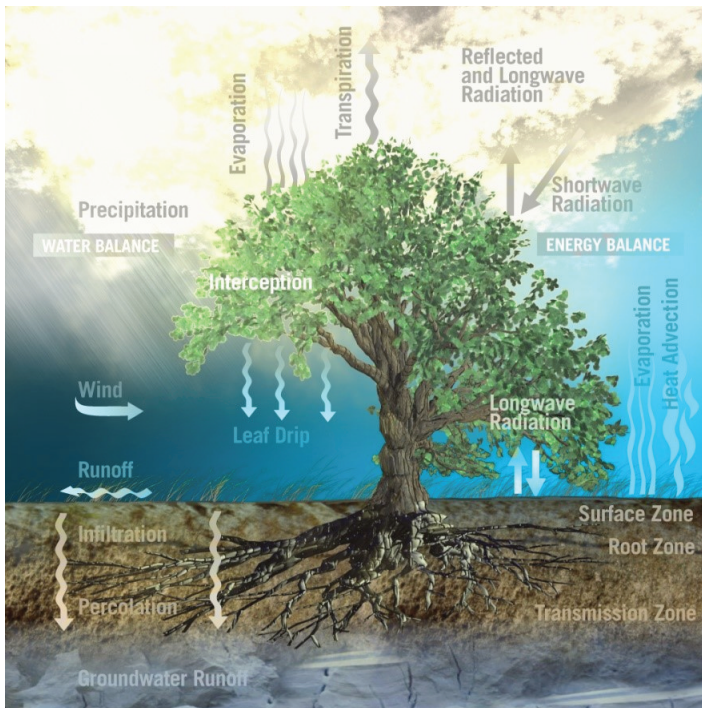
Fig. 7-1. The processes of the soil water balance (left side) and the land surface energy balance (right side). Source: European Space Agency (link).

At the outset, it may be helpful to consider the value question: is runoff a good thing or a bad thing? The answer depends on the context. For example, runoff from cropland is typically undesirable because it results in reduced water availability for the crop, erosion of fertile topsoil, and pollution of surface water bodies with sediment, phosphorus, and other contaminants (Fig. 7-2). In contrast, runoff from uncultivated or impervious areas can sometimes be beneficial when captured and stored for later use. Such water harvesting strategies may be key solutions to meeting critical water challenges in developing regions such as sub-Saharan Africa [1] and even in major metropolitan areas like Sydney, Australia, located in arid or semi-arid regions of developed nations [2]. Runoff is also one of the main sources of water to surface water bodies, so any soil or water management practices that influence runoff are likely to influence surface water quantity and quality.

Fig. 7-2. *Runoff and water erosion from this corn (maize) field in Iowa, USA, has harmful onsite and offsite effects (left). The water in this runoff-harvesting pit in Uganda is used to irrigate bananas, cassava, corn, and vegetables (right). Photo credits: Lynn Betts, USDA NRCS (link) and Matilda Nakawungu, Global Water Initiative – East Africa (link).*

7.1 Overview of Runoff and Water Erosion Processes

Before runoff (sometimes called **overland flow**) begins, there must be excess water on the soil surface. Conceptually, that excess water can occur for one of two reasons:

- the soil is unsaturated but the rate of water delivery to the surface exceeds the infiltration rate into the soil, this is called **infiltration-excess overland flow**, or
- the soil is saturated (or satiated) and cannot accept any more water, this is called **saturation-excess overland flow**.

In either case, runoff does not begin immediately upon the generation of the first excess water. The soil surface is never perfectly smooth, and the soil roughness results in a finite **surface storage capacity**, which is the volume of excess water per unit area that can be retained on the surface. Runoff can only begin once the surface storage capacity is exceeded.

When runoff occurs, the soil surface is susceptible to water erosion, i.e. erosion due to flowing water. *The erosion process includes three primary stages: detachment, transport, and deposition.* At the soil surface, detachment of soil particles from the bulk of the soil body can occur due to raindrop impact, the breakdown of soil aggregates upon wetting, and the scouring force of surface runoff. The rate of detachment during water erosion depends on:

- the degree to which the soil surface is covered by vegetation, plant residues, or other protective covers,
- the soil strength (recall Fig. 5-9),
- the rainfall intensity,
- and the velocity of the surface runoff.

The transport of water and sediment, i.e. detached soil particles, across the soil surface during a runoff event often follows distinct patterns, which represent different types of water erosion. The first type is **sheet erosion**, which is characterized by water flow and soil erosion distributed relatively uniformly across the soil surface. This type of erosion can be particularly insidious, sometimes going relatively unnoticed for years and causing major soil degradation. The images in Fig. 7-3 show an example where >30 cm of topsoil were lost to sheet erosion, effectively destroying the productive capacity of the soil.
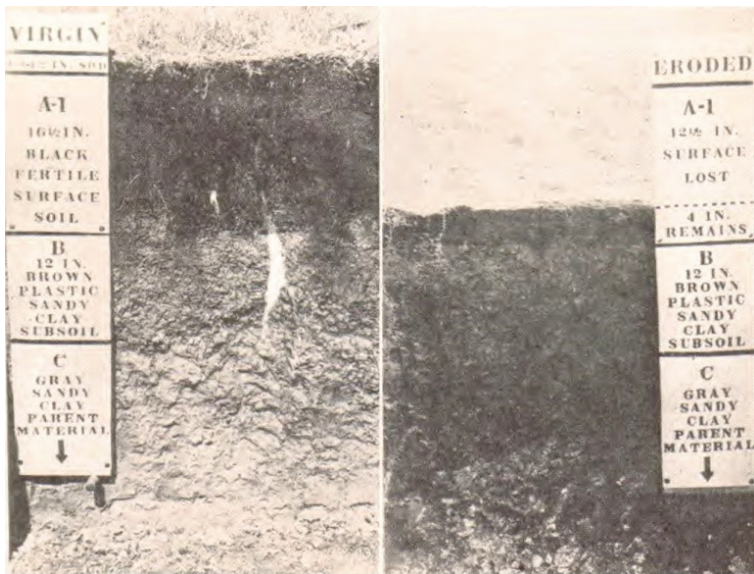
Fig. 7-3. Comparison of two fields in Missouri with the same soil type, showing the amount of topsoil lost by sheet erosion. Reproduced from: Soil Erosion and Its Control, Q.C. Ayers, 1936, McGraw Hill.

When water and sediment flowing across the surface first begin to concentrate in small, shallow channels, the result is **rill erosion**. As the water continues to flow downhill, rills may deepen and widen

to form gullies, resulting in **gully erosion**. Rills are distinguished from gullies in that rills can often be removed using tillage and are easily crossed with field equipment, while neither of these is true for gullies. The distinction between rill and gully erosion is illustrated in Fig. 7-4.

Fig. 7-4. Rill erosion on highly erodible soil in northwest Iowa after heavy spring rains (left, Source: Lynn Betts, USDA NRCS, link) and severe gully erosion on loess soils in western Iowa (right, Source: USDA NRCS, link).

Other types of water-related erosion are driven by the force of gravity acting on wet soil along hillslopes, gullies, and streambanks. As soil in these locations becomes thoroughly wetted, the weight of the soil body increases and the soil strength decreases raising the risk of gradual soil creep, sudden slumps, and potentially devastating landslides (Fig. 7-5). On January 13, 2001, a 7.6 magnitude earthquake struck El Salvador, setting off a massive landslide that killed ~585 people in the residential area of Santa Tecla. Wet soil is thought to have been a key factor contributing to this tragedy.

Fig. 7-5. Soil creep (left) and soil slump (center) in North Dakota, USA. Deadly landslide in Santa Tecla, El Salvador. Sources: North Dakota State University (link) and USGS (link).

The final stage of the erosion process is sediment deposition. Sediment deposition is a major issue affecting streams, reservoirs, and coastal areas. It is also one of the primary drivers of the dramatic spatial variability found in alluvial soils. Sediment control is an important management concern in agriculture, construction, and engineering. Sediment deposition is typically initiated by a decrease in the flow velocity and can be approximated using Stokes' Law (Eq. 2-2).

The three stages of water erosion as a function of flow velocity are depicted for different particle sizes in a helpful diagram called the Hjulström-Sundborg diagram (Fig. 7-6). Based on work of researchers at the University of Uppsala in Sweden, this diagram summarizes the behavior of particles in a stream or river [3]. Erosion in this context means the detachment of a particle from the streambed. Note that both axes have a logarithmic scale. Several interesting features of erosion, transport, and deposition processes can be seen by carefully inspecting the diagram.
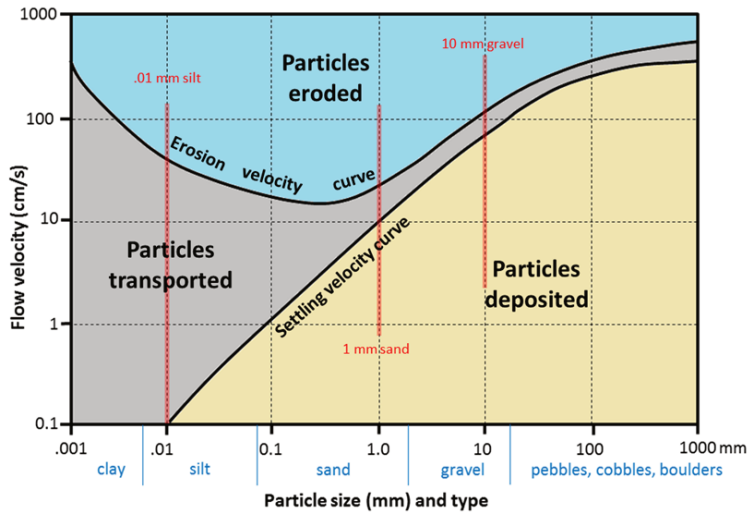
Fig. 7-6. The Hjulström–Sundborg diagram showing the relationships between particle size and the tendency to be eroded, transported, or deposited at different current velocities. By Steven Earle, CC-BY 4.0 (link).

For example, we can see that a water flow velocity of roughly 20 cm s^{-1} is required to detach a 1-mm sand-sized particle from the streambed [4]. However, once that particle is detached, it will remain in suspension and be transported by the flowing water until the flow velocity drops below 10 cm s^{-1}. The diagram also shows that sand size particles in the range from approximately 0.2-0.5 mm (fine to medium sand) have the lowest detachment velocity and are thus most erodible. The detachment velocity increases for smaller particle sizes, such as silt and clay, because the attractive and adhesive forces between the particles increase as the particle size decreases. Finally, the diagram highlights the fact that clay size particles, once detached and suspended, will not be deposited but will remain in suspension indefinitely even at flow velocities as low as 0.1 cm s^{-1}.

7.2 Runoff and Erosion Models

Over the years, scientists and engineers have developed many models for predicting runoff and erosion, and these models range from relatively simple empirical models to highly complex mechanistic models. One of the most widely used empirical models for predicting runoff is the ***Curve Number Method***, which was developed by the USDA Soil Conservation Service (now the Natural Resources Conservation Service, NRCS) in the 1950's [5]. The curve number method is based on the idea of a ratio between the actual amount of precipitation retained by the landscape during a rainfall event, F, and the potential maximum retention amount for that landscape, S. The *core hypothesis* of the method is that this ratio (F/S) is equal to the ratio between the amount of runoff from the landscape, Q, and the total precipitation amount, P, reduced by some "initial abstraction", I_a. Mathematically, we can write this as:

$$[\text{MISSING EQUATION}] \quad \text{(Eq. 7-1)}$$

Now, the amount retained during the event, F, is equal to the total rainfall minus the initial abstraction minus the amount of runoff, i.e. **[MISSING EQUATION]**. Substituting this expression for F in Eq. 7-1 and rearranging gives:

$$[\text{MISSING EQUATION}] \quad \text{(Eq. 7-2)}$$

which is only valid for **[MISSING EQUATION]**. The initial abstraction is typically estimated as 20% of the potential maximum retention, i.e. **[MISSING EQUATION]**. The value of S is calculated based on an empirical ***curve number***, CN, which is a number between 0 and 100 selected based on the hydrologic characteristics of the landscape.

$$[\text{MISSING EQUATION}] \quad \text{(Eq. 7-3)}$$

where S here is in inches. Unit conversion can be applied as needed, and the units of S, P, I_a, and Q must be the same. Curve

numbers for a wide variety of circumstances are available here (link). The expected relationship between rainfall and runoff based on the curve number method is shown in Fig. 7-7 for a range of CN values. Procedures exist for adjusting the CN values based on the initial soil moisture conditions, with lower values for drier conditions and higher values for wetter conditions. Research has indicated that setting the initial abstraction to 5%, rather than 20%, of S results in more accurate runoff predictions, but this change also requires revising the existing CN values [6].
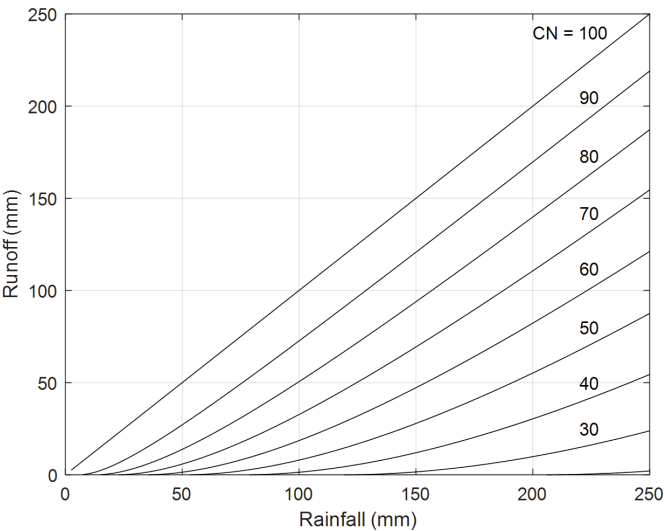
Fig. 7-7. *Predicted runoff depth (mm) as a function of rainfall depth (mm) for a range of curve numbers (CN) based on the assumption that $I_a = 0.2S$.*

In contrast to empirical methods like the Curve Number Method, mechanistic models have been developed, which predict runoff by simulating the underlying mechanisms of the soil water balance and the overland flow processes. These models typically require detailed information about weather and rainfall patterns, soil properties,

topography, and land use. This information is used to simulate the dynamics of rainfall interception and infiltration, and from those, runoff is predicted. Many of these models include spatial information allowing simulation of overland flow processes as runoff from different portions of the landscape converges towards a stream. One widely used mechanistic model capable of predicting both runoff and the associated soil erosion is the USDA Water Erosion Prediction Project (WEPP) model.

The WEPP model development was led by the National Soil Erosion Research Laboratory of the USDA Agricultural Research Service. Over 200 people contributed to the development of the model, and the team conducted rainfall simulation experiments (Fig. 7-8) at over 50 sites to generate data for developing and testing the model [7]. The model simulates infiltration using the Green-Ampt approach in a manner similar to that described in section 6.3. As we saw previously, the soil hydraulic conductivity is a key parameter in the Green-Ampt approach, and in the WEPP model the hydraulic conductivity is influenced by tillage, crusting, surface cover, and storm precipitation amount [8]. Runoff is predicted based on the surface water excess. Despite its complexity, the WEPP model can be setup and run relatively easily using the online versions of the model (link).
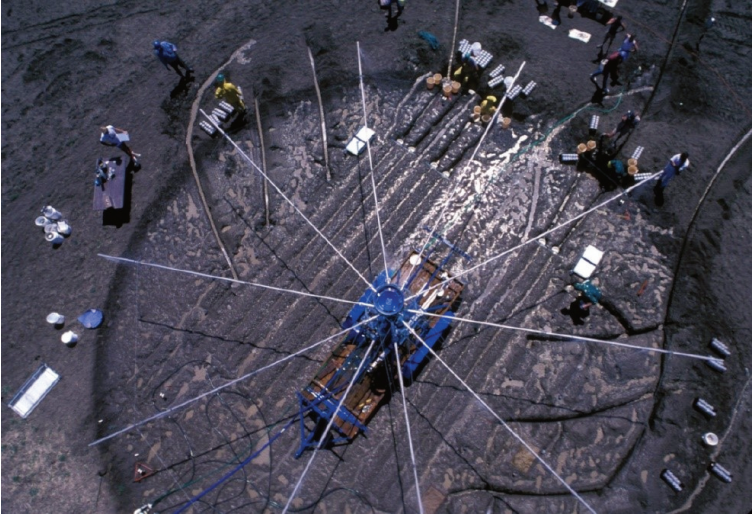
Fig. 7-8. Overhead photograph of a rainfall simulation experiment near Cottonwood, South Dakota. This is one of many experiments conducted to support the development of the WEPP model. Reproduced from Flanagan et al. (2007). Photo source: USDA ARS, Photographer: Tim McCabe.

7.3 Problem Set

1. Use the curve number (CN) method to estimate the total runoff (mm) and the percent of annual precipitation partitioned to runoff (%) during the year 2013 for a field near Stillwater, OK, under two different scenarios:

   ◦ Fallow – bare soil with poor hydrologic condition, hydrologic group C, and CN = 91.
   ◦ Small grain – contour farmed with adequate crop residue, soil with poor hydrologic condition, hydrologic group C, and CN = 81.

   An Excel spreadsheet with 2013 daily rainfall data from the Oklahoma Mesonet is available online at: http://soilphysics.okstate.edu/teaching/soil-4683/homework/

     Work using rainfall in inches, and convert the final result to mm. Perform daily runoff calculations, then compute the annual sum. The "IF" function in Excel may be helpful.

2. Use the Online GIS Interface for the WEPP model (link) to estimate runoff and erosion for a watershed of your own choosing following these steps:

   a. Zoom in to an area in the United States covering ~1 square mile.
   b. Build the channel network.
   c. Select the watershed outlet point of interest.
   d. Build the subcatchments.
   e. Review the watershed summary, *recording*:

      ▪ outlet coordinates to three decimal places
      ▪ two largest land uses and their percentages

- two largest soil types and their percentages

  f.   Close the watershed summary window.

  g. If necessary, change land use associations.

  h. Setup the WEPP model.

  i.   ***Record*** the distance to the selected climate station.

  j. Specify watershed and flow path simulation for a 10 year period.

  k. Instruct the model to determine land use and soil type for each grid cell.

  l. Run the WEPP model. This may take several minutes.

 m. After the simulations are completed, view the erosion maps.

  n. View the simulation summary, ***recording***:

- the watershed area (ha)
- the annual precipitation (mm yr-1)
- the annual discharge (mm yr-1)
- the sediment yield (t ha-1 yr-1)

  o.   If desired, save your project for future reference.

7.4 References

1. Dile, Y.T., et al., *The role of water harvesting to achieve sustainable agricultural intensification and resilience against water related shocks in sub-Saharan Africa.* Agriculture, Ecosystems & Environment, 2013. **181**: p. 69-79.
2. Rahman, A., J. Keane, and M.A. Imteaz, *Rainwater harvesting in Greater Sydney: Water savings, reliability and economic benefits.* Resources, Conservation and Recycling, 2012. **61**: p. 16-21.
3. Sundborg, Å., *Some aspects on fluvial sediments and fluvial morphology I. General views and graphic methods.* Geografiska Annaler: Series A, Physical Geography, 1967. **49**(2-4): p. 333-343.
4. Earle, S., *Physical Geology.* BC Open Textbook Project. 2016.
5. Boughton, W., *A review of the USDA SCS curve number method.* Soil Research, 1989. **27**(3): p. 511-523.
6. Woodward, D.E., et al., *Runoff curve number method: examination of the initial abstraction ratio*, in *World Water & Environmental Resources Congress 2003.* 2003: Philadelphia, Pennsylvania. p. 1-10.
7. Flanagan, D.C., J.E. Gilley, and T.G. Franti, *Water Erosion Prediction Project (WEPP): Development history, model capabilities, and future enhancements.* Transactions of the ASABE, 2007. **50**(5): p. 1603-1612.
8. Flanagan, D. and M. Nearing, *USDA-Water Erosion Prediction Project: Hillslope profile and watershed model documentation.* 1995, NSERL report.

8. REDISTRIBUTION AND DRAINAGE

We have considered how water inputs are modified by the process of interception and then partitioned between infiltration and runoff. Now, we will turn our attention to the fate of the infiltrated water. The process of continued movement of soil water after infiltration ends is called *redistribution*. In this chapter, we will focus on some fundamental aspects of the redistribution process such as:

- its rate decreases over time
- it influences the availability of water to plants
- it drives solute transport in soil

We will also examine the process of *drainage*, which is the loss of water from below the root zone or from the bottom of the soil profile. We will learn a relatively simple method for estimating drainage rates and will also think about artificial (or man-made) soil drainage systems and their impact on agriculture and the environment.

8.1 Redistribution Within Partially Wetted Profiles

At the end of an infiltration event, the soil profile is often in a condition far from hydraulic equilibrium. Water continues to flow from the regions of higher water potential, that is the portion of the soil wetted during infiltration, to regions of lower water potential located farther from the infiltrating surface. Measured water content profiles during two different redistribution events for a clay soil are shown in Fig. 8-1 [1]. The top portion of the figure shows redistribution following a 1.5 inch (3.8 cm) infiltration event. The wetting front had reached a depth of ~10 cm at the end of the infiltration event. During redistribution, water flowed out of the soil above that depth, and the *water content decreased* from ~0.44 cm^3 cm^{-3} at the start of redistribution to ~0.32 cm^3 cm^{-3} after one day. The water which flowed out of the infiltration–wetted portion of the soil flowed into the soil beneath, such that at the 12-cm depth, the *water content increased* from ~0.04 cm^3 cm^{-3} at the end of infiltration to ~0.28 cm^3 cm^{-3} after one day of redistribution. There was also a thin layer of soil centered on the 11-cm depth in which the water content increased after one hour of redistribution and then decreased after one day.
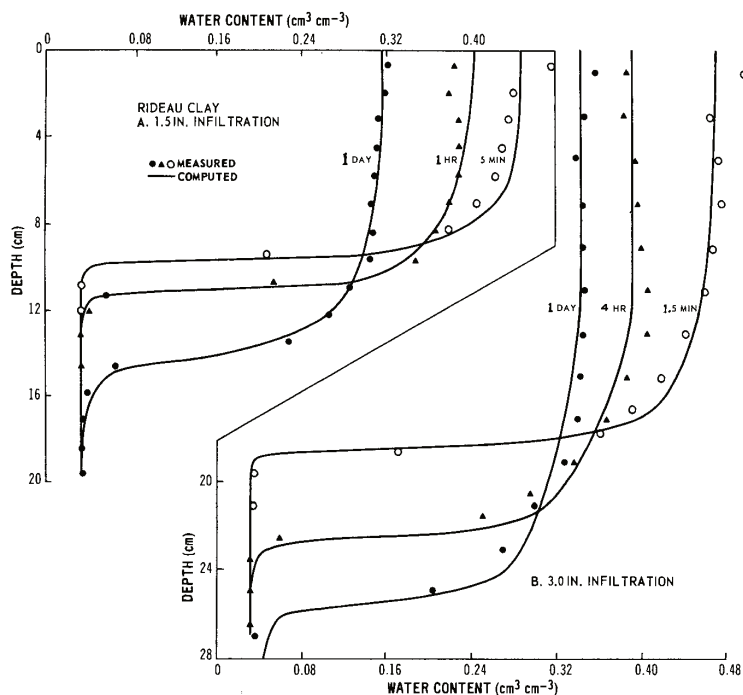
Fig. 8-1. Soil water profiles during redistribution for a clay soil after 1.5 inches (3.8 cm) of infiltration (A) and after 3.0 inches (7.6 cm) of infiltration (B). Reproduced from Staple (1969).

The data in Fig. 8-1 help highlight some other key features of the redistribution process. You can see that the water content gradient associated with the wetting front diminishes over time during redistribution. Likewise, the *matric potential gradient associated with the wetting front also diminishes over time*, and this causes the rate of redistribution to decrease over time. Furthermore, as the water content decreases in the infiltration-wetted portion of the profile, *the hydraulic conductivity of that soil decreases dramatically*, and this also causes the rate of redistribution to decrease over time.
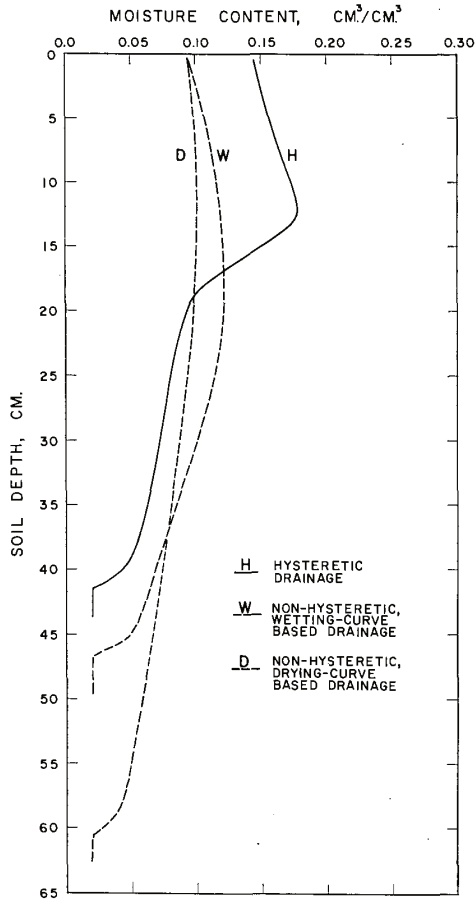
Fig. 8-2. Hysteretic and non-hysteretic water content profiles for a sand column after 2 hours of redistribution. Reproduced from Rubin (1967).

The process of soil water redistribution creates a dynamic situation where the infiltration-wetted part of the soil profile is getting drier while the neighboring parts of the soil are getting wetter. For this reason, the redistribution process is sensitive to hysteresis in the water retention curve. As you may recall from Section 3.3.3, the

water content for any given matric potential is higher for the drying branch of the water retention curve than for the wetting branch (Fig. 3-8). As a result, the infiltration-wetted part of the soil, the part which is getting drier during redistribution, *retains more water than it would if hysteresis did not occur.* This phenomenon is evident in the water content profiles in Fig. 8-2, which were calculated using a numerical model of the redistribution process [2]. Thus, there are three mechanisms that may "help" the soil retain water near the soil surface after infiltration of rainfall or surface irrigation:

- the decreasing matric potential gradient,
- the decreasing hydraulic conductivity in the infiltration-wetted zone, and
- the effects of hysteresis.

8.2 Drainage from the Soil Profile

The strong forcings imposed at the soil surface by rainfall or irrigation events and by diurnal and annual cycles of solar radiation are typically moderated as we move deeper in the soil profile. As a result, the gradients in pressure potential are often smaller in magnitude at depth than they are near the soil surface. At sufficient depth, the pressure potential gradient may become negligible and the drainage rate can be approximated using the **unit-gradient approach**.

That approach begins with the Buckingham-Darcy Law

$$q = -K(\Theta)\frac{d(\Psi_g + \Psi_p)}{dz} = -K(\Theta)\left[\frac{d\Psi_g}{dz} + \frac{d\Psi_p}{dz}\right]$$

(Eq. 8-1)

where q is the water flux, which in this case is the drainage rate; $K(\theta)$ is the hydraulic conductivity, which is a function of the volumetric water content, θ; d/dz indicates a gradient in the vertical direction, z, which is defined here as positive downward; ψ_p is the gravitational potential; and yp is the pressure potential. Assuming that the pressure potential gradient, the last term in Eq. 8-1, is negligible, and noting that $d\psi_g/dz$ = -1 when z is positive downward, we can simplify the Buckingham-Darcy Law in this case as:

$$q = K(\Theta)$$

(Eq. 8-2)

This unit-gradient equation tells us that, when the pressure potential gradient is negligible at a specific depth in the soil profile, the downward drainage rate at that depth is equal to the hydraulic conductivity value associated with the soil water content at that

depth. Here it is important to remember that the soil hydraulic conductivity has a strong, non-linear relationship to soil water content (recall Fig. 4-7). Therefore, accurate drainage estimates from the unit-gradient approach require careful determination of that relationship.
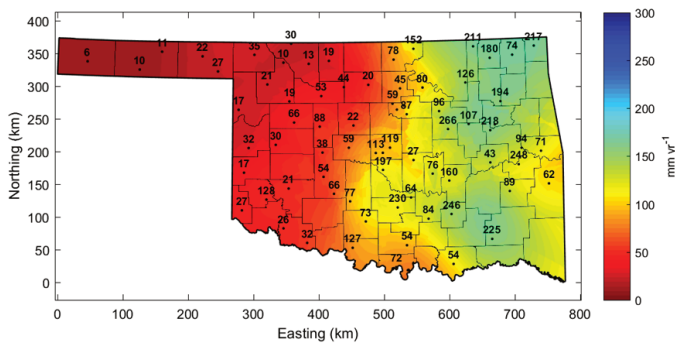
Fig. 8-3. Mean annual drainage rates (mm yr^{-1}) across the state of Oklahoma for the years 1998-2014. Drainage was estimated by the unit-gradient approach using daily soil moisture data at the 60-cm depth from the stations of the Oklahoma Mesonet. The station locations are shown by the black dots. Drainage rate labels for the Stillwater, Oklahoma City East, Porter, and Marena sites were excluded for clarity, but were 214, 82, 166, and 66 mm yr^{-1}, respectively.

The unit-gradient approach has sometimes been employed to estimate groundwater recharge, on the assumption that water draining below the chosen measurement depth will eventually reach the groundwater table [3]. In a recent study in Oklahoma, the unit-gradient approach was applied to soil moisture data from a statewide monitoring network called the Oklahoma Mesonet [4].
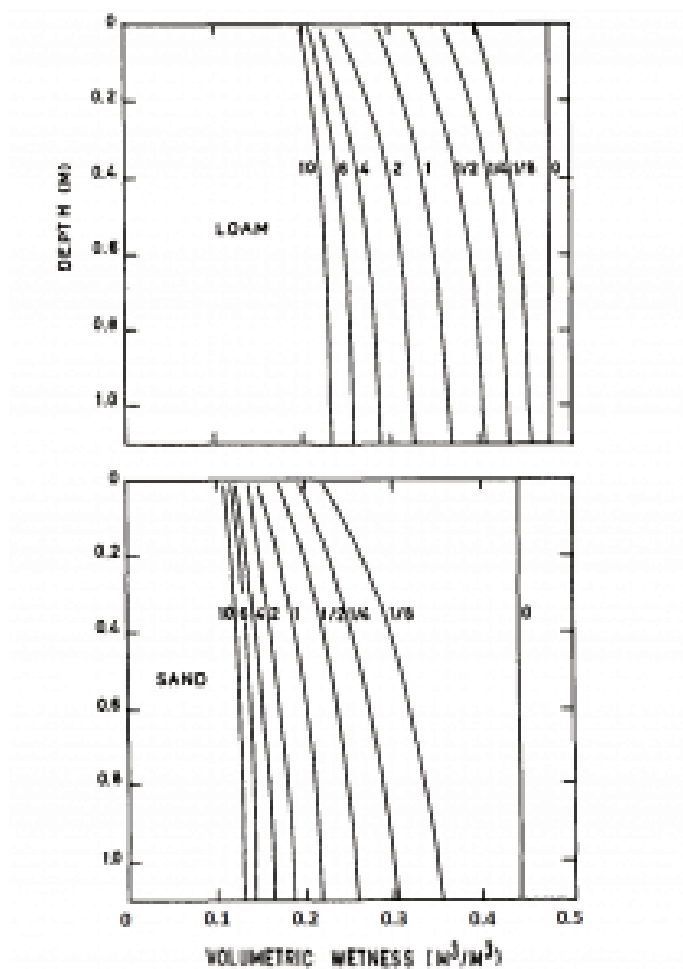
Fig. 8-4. Simulated soil moisture distributions during drainage from initially-saturated uniform profiles of loam (upper) and sand (lower). The numbers indicate duration of the process (days). Adapted from Hillel and van Bavel (1976).

Daily soil moisture data for the 60-cm soil depth were used in

the Mualem-van Genuchten hydraulic conductivity function (Eq. 4-8). The soil parameters for that function were estimated using a pedotransfer function called Rosetta [5]. The resulting drainage estimates (Fig. 8-3) reflect the gradient in annual precipitation, which ranges from approximately 1420 mm in the southeast part of the state to 430 mm in the western Oklahoma Panhandle. Annual drainage maps for specific years show large year-to-year fluctuations driven by differences in annual precipitation (link).

The rate and cumulative amount of drainage is influenced, not only by precipitation, but also by the soil texture. All other factors being equal, *fine-textured soils exhibit lower drainage rates and less cumulative drainage than do coarse-textured soils.* This is a direct result of the larger pore sizes typical in coarse-textured soils, leading to larger values for hydraulic conductivity at and near saturation, compared to those in fine-textured soils. Numerical simulations of the drainage process for uniform profiles of loam and sand (Fig. 8-4) illustrate the dramatic effects of soil texture on drainage [6]. During the first day of drainage, the volumetric water content of the sand decreased from approximately 0.45 cm^3 cm^{-3} to approximately 0.20 cm^3 cm^{-3}. During the same time interval, the loam exhibited much less drainage with the volumetric water content decreasing from approximately 0.47 cm^3 cm^{-3} to approximately 0.35 cm^3 cm^{-3}. Even less drainage would occur for a clay soil under the same conditions.

In addition to the effects of soil texture, another key feature of the drainage process is evident in Fig. 8-4. Notice that the drainage rates decreased over time, as indicated by the progressively smaller distances between the water content curves. *However, there is no distinct time at which drainage ceased.* Even after 6 days of drainage in a uniform sand, the water content continued to decrease and the drainage process continued. The fact that the drainage process typically does not have a distinct endpoint highlights a major problem with a widely-used concept in soil science, agronomy, and hydrology–the concept of "field capacity".

8.3 Field Capacity

The concept of field capacity has often been used in irrigation planning, crop modeling, and hydrologic modeling. **Field capacity** is thought to be the volumetric water content at which drainage effectively ceases. This hypothetical water content value is often incorrectly considered a property of the soil. However, the whole field capacity concept is flawed because, as we have just seen, there are typically no clearly defined breakpoints at which drainage stops. For example, in a drainage experiment on a silt loam soil in Israel, the gravimetric water content in the 60- to 90-cm depth decreased from 0.29 g g^{-1} at the start of drainage, to 0.20 g g^{-1} after 1 day, 0.19 g g^{-1} after 2 days, 0.16 g g^{-1} after 30 days, and 0.15 g g^{-1} after 60 days [7]. At which of these water contents could we accurately say that drainage has ceased?

Procedures have been developed for estimating field capacity in the field and in the laboratory. The most common way to estimate field capacity is to assume that it is equal to the water content retained in the soil at a specific matric potential. Research has proven repeatedly that there is not any one matric potential value which universally represents field capacity, but the convenience of this approach to estimating field capacity leads to its continued use. In the United States, a matric potential of -33 kPa has most commonly been used for this purpose. In some other countries, a value of -10 kPa has been more widely-used. Recent research has shown that -33 *kPa provides a poor approximation of field capacity in most cases*, and less negative values of matric potential such as -10 or even -6 kPa appear to be more suitable [8, 9]. Alternative approaches for estimating field capacity based on the parameters of the soil water retention curve have also been proposed [10, 11].

8.4 Artificial Drainage

In some situations, the natural drainage capability of the soil is inadequate to meet the requirements of the intended agricultural or engineering uses for the land. Therefore, people invest substantial time, money, and resources to increase soil drainage rates by installing artificial drainage systems. One widespread type of artificial drainage is subsurface drainage, which in the US is commonly called **tile drainage**. That name derives from the clay tiles that were used to construct these underground drainage systems in the past. Now, subsurface drainage systems are constructed mainly from heavy-duty, corrugated, perforated plastic tubing. This tubing if often installed approximately 1.2 m below-ground using a large tractor and a special implement called a tile plow (Fig. 8-5).

Fig. 8-5. Tractor and tile plow installing plastic drainage tile in cropland. Source: Ag Leader Technology (link).

Subsurface drainage systems can remove not only excess water but also nutrients and pesticides from the soil. These compounds can negatively affect the water quality and aquatic ecosystems of the surface water bodies that receive the drainage water. One large-scale example is the **hypoxic zone** that occurs where the Mississippi River flows into the Gulf of Mexico. In this context, the hypoxic zone is defined as the region where the concentration of dissolved oxygen in the water at the sea floor is <2 mg L^{-1}. At these low oxygen levels bottom-dwelling fish and shrimp cannot survive, so the hypoxic zone has also been called the "dead zone".

Subsurface drainage systems have substantially increased the agricultural productivity of poorly-drained soils around the world. The increased drainage improves the soil's ability to transport oxygen to crop roots to support cellular respiration. This allows crops to develop deeper and larger root systems, which enable higher yields. These agronomic benefits have driven broad adoption of this technology in areas with poorly-drained soils. The map in Fig. 8-6 shows the percent of the total land area in each US county estimated to be artificially drained cropland [12]. Artificial drainage is estimated to affect >10% of the land area in many counties and >50% in some portions of the Mississippi River floodplain and the Midwest. Unfortunately, the agronomic benefits of these drainage systems often come with unintended ecological and environmental consequences.
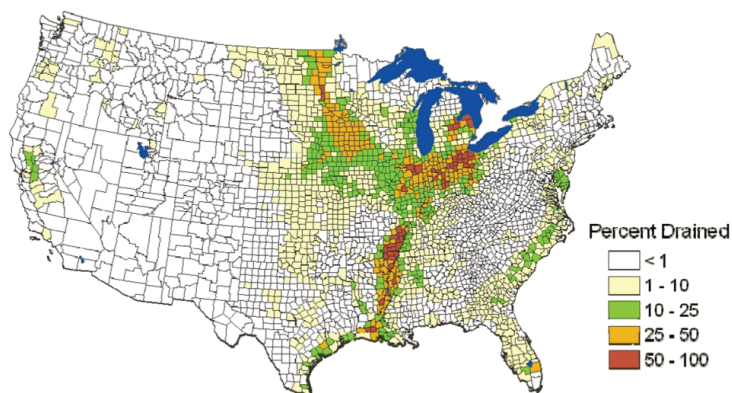
Fig. 8-6. Percent of the total land area in each county estimated to be artificially drained cropland. Reproduced from Jaynes and James (2007).

Hypoxia develops when nutrient rich water flows into the Gulf enhancing the growth of algae and other organisms near the surface. When these organisms die and settle to the sea floor, bacteria decompose the dead organisms, consuming oxygen in the process resulting in hypoxia. The loss of nutrients, particularly nitrogen and phosphorus, from cropland in the Mississippi River basin is considered to be a major contributor to the growth in size of the hypoxic zone in the Gulf of Mexico since the 1980s (Fig. 8-7). And, research has shown that artificial drainage systems have increased nutrient loss from cropland, making artificial drainage one contributing factor to this large-scale environmental problem.

2014 DEAD ZONE Area of oxygen-deprived bottom-water

Mississippi River

TEXAS   LOUISIANA   Lake Pont.   MISS.

Baton Rouge

Lake Charles   Lafayette

New Orleans

Houma

Gulf of Mexico

50 miles

5,052 square miles

Normal oxygen levels   Lowest oxygen levels

KEY:

Sources: Nancy Rabalais, Louisiana Universities Marine Consortium; Eugene Turner, Louisiana State University; NOAA Center for Sponsored Coastal Ocean Research

NOLA.com | The Times-Picayune

Fig. 8-7. Map of the Gulf of Mexico hypoxic zone or "dead zone" in 2014.

Recognizing the potential environmental impacts of drainage is important, but it is also important to understand that drainage itself is not necessarily problematic. Redistribution of water within the soil profile and drainage of water from the soil profile are essential to the functioning of many natural and managed landscapes. These processes can help sustain plant growth, biogeochemical cycling, and groundwater aquifers. Redistribution and drainage also strongly influence the movement of solutes in the environment, and in the next chapter we will consider solute transport processes in more detail.

8.5 Problem Set

1. Use the unit-gradient approach and Campbell's hydraulic conductivity model (Eq. 4-7) to:

   a. estimate the drainage rate for a loam soil with K_s = 310 mm d^{-1}, b = 4.5, θ_s = 0.46, and θ = 0.40, and

   b. estimate the volumetric water content at which the drainage rate would be 1 mm d^{-1}.

2. A uniform silt loam soil is initially wetted to θ = 0.42 cm^3 cm^{-3} to a depth of 1.2 m. The soil has a K_s value of 160 mm d^{-1}, θ_s = 0.44 cm^3 cm^{-3}, and b = 4.7 in the Campbell hydraulic conductivity model.

   a. What is the value of θ when the drainage rate drops to 1 mm d^{-1}, assuming unit-gradient conditions?

   b. How many mm of water have drained from the soil profile when that value of θ is reached?

8.6 References

1. Staple, W.J., *Comparison of Computed and Measured Moisture Redistribution Following Infiltration1*. Soil Science Society of America Journal, 1969. **33**(6): p. 840-847.

2. Rubin, J., *Numerical Method for Analyzing Hysteresis-Affected, Post-Infiltration Redistribution of Soil Moisture1*. Soil Science Society of America Journal, 1967. **31**(1): p. 13-20.

3. Nolan, B.T., et al., *Factors influencing ground-water recharge in the eastern United States*. Journal of Hydrology, 2007. **332**(1–2): p. 187-205.

4. McPherson, R.A., et al., *Statewide Monitoring of the Mesoscale Environment: A Technical Update on the Oklahoma Mesonet*. Journal of Atmospheric and Oceanic Technology, 2007. **24**(3): p. 301-321.

5. Schaap, M.G., F.J. Leij, and M.T. van Genuchten, ROSETTA: *a computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions*. Journal of Hydrology, 2001. **251**(3-4): p. 163-176.

6. Hillel, D. and C.H.M. van Bavel, *Simulation of Profile Water Storage as Related to Soil Hydraulic Properties*. Soil Science Society of America Journal, 1976. **40**(6): p. 807-815.

7. Hillel, D., *Soil and Water: Physical Principles and Processes. .* 1971, New York: Academic Press. 288.

8. de Jong van Lier, Q. and O. Wendroth, *Reexamination of the Field Capacity Concept in a Brazilian Oxisol*. Soil Science Society of America Journal, 2016. **80**(2): p. 264-274.

9. Ottoni Filho, T.B., et al., *An Algebraic Pedotransfer Function to Calculate Standardized in situ Determined Field Capacity*. 2016, 2016. **8**(8).

10. Assouline, S. and D. Or, *The concept of field capacity revisited: Defining intrinsic static and dynamic criteria for soil internal drainage dynamics*. Water Resources Research, 2014. **50**(6): p.

4787-4802.

11. Twarakavi, N.K.C., M. Sakai, and J. Šimůnek, *An objective analysis of the dynamic nature of field capacity.* Water Resources Research, 2009. **45**(10).

12. Jaynes, D.B. and D.E. James, *The Extent of Farm Drainage in the United States.* 2007, USDA Agricultural Research Service, National Soil Tilth Laboratory: Ames, Iowa.

9. SOLUTE TRANSPORT AND GROUNDWATER POLLUTION

In the prior chapter, we considered drainage from the soil profile and one example of solute transport. In this chapter, we will focus more directly on solute transport and the related issues of groundwater pollution and soil salinity. Understanding the mechanisms and pathways of solute transport is important for addressing an array of pressing environmental problems from drinking water contamination to land degradation. The mechanisms of solute transport also strongly influence plant nutrient uptake.

9.1 Solute Transport

The primary mechanisms for solute transport in soil are advection and diffusion. **Advection** is the transport of solutes due to the bulk flow of water. **Diffusion** is the net movement of solutes from a region of higher concentration to a region of lower concentration due to the random thermal motion of the solute and water molecules. We will begin our study of solute transport by focusing on these two fundamental mechanisms.

9.1.1 Advection

When water flows through soil with a flux q and a solute concentration C, the resulting advective flux of the solute, J_a, is:

$$J_a = q \cdot C$$

<div align="right">(Eq. 9-1)</div>

and has units of mass of solute passing through a unit cross-sectional area of soil per unit time (e.g. $g\,m^{-2}\,yr^{-1}$). When estimating the distance a solute may travel in a certain amount of time, we need to consider not only the water flux, q, but also the pore water velocity, v. The **pore water velocity** is the average distance in the direction of the bulk flow which is traveled by an individual water molecule in a unit of time. The pore water velocity depends on the flux and the volumetric water content, θ:

$$\nu = \frac{q}{\Theta}$$

<div align="right">(Eq. 9-2)</div>

Thus, for a given flux, the pore water velocity is higher when the volumetric water content is lower. If we could measure the water velocities in the soil at the pore scale, we would find a complex

pattern of flow paths heading in different directions and at different speeds, as suggested by the simulation results shown in Fig. 9-1. The average pore water velocity, v, is the net outcome of these innumerable complex flow paths.
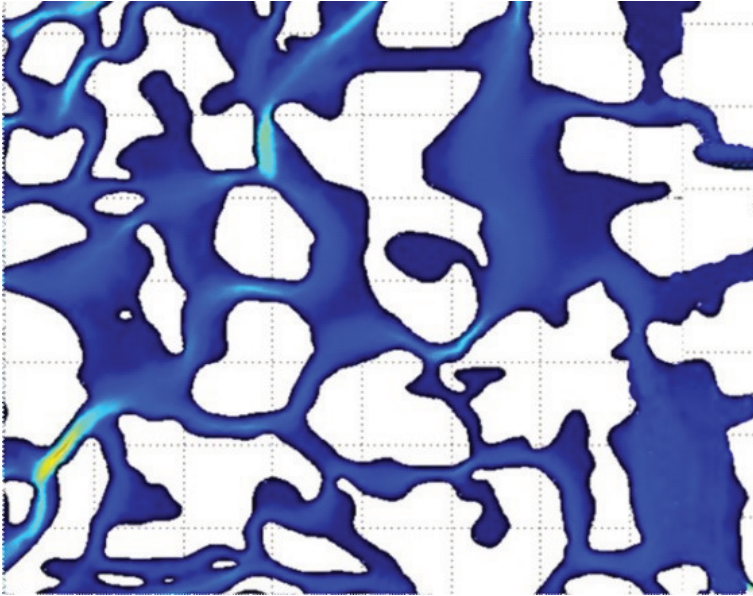
Fig. 9-1. Heat map of simulated pore-scale water velocities. The white spaces represent the 2-D porous media. The brightly colored areas have relatively high velocities while the dark blue areas have relatively low velocities. Source: http://www2.bren.ucsb.edu/~keller/micromodels.htm

Once we know the pore water velocity, we can make a rough estimate of the travel time, t, required for a solute to pass through a layer of soil of specified thickness, L:

$$t = \frac{L}{\nu}$$

(Eq. 9-3)

If the soil layer of interest has a uniform water content, we can combine Eqs. 9-3 and 9-2 to get:

$$t = \frac{L\Theta}{q}$$

These transport time estimates are only useful for solutes that do not interact in any way with the soil. Adjustments must be made for solutes that interact with the soil, a case we will consider in a later section.

9.1.2 Diffusion

Solute **diffusion** is the net movement of solutes from a region of higher concentration to a region of lower concentration due to the random thermal motion of the solute and water molecules. Mathematically, this diffusion is governed by Fick's Law:

$$J_d = D\frac{dC}{dZ}$$

where D is the diffusion coefficient for a particular solute and dC/dz represents the gradient in the solute concentration. The diffusion coefficient for a solute in soil depends on:

- the diffusion coefficient of that solute in pure water,
- the volumetric water content of the soil, and
- the tortuosity of the diffusion paths within the soil pore network.

Solute diffusion in soil is hindered as the soil water content decreases and the tortuosity of the diffusion pathways increases. Thus, diffusion is often a relatively minor contributor to solute transport in soil. However, diffusion can be an important

mechanism of transport over small distances, as in the case of nutrient uptake by plant roots.

9.1.3 Hydrodynamic Dispersion

Under the correct circumstances, advection and diffusion interact in a unique way to produce a composite phenomenon called hydrodynamic dispersion. **Hydrodynamic dispersion** is the process in which the solute appears to be diffusing uniformly upstream and downstream from a plane which is moving downstream at a rate equal to the average pore water velocity. The simulations shown in Fig. 9-2 show how this dispersion process arises for the special case of flow through a capillary tube.
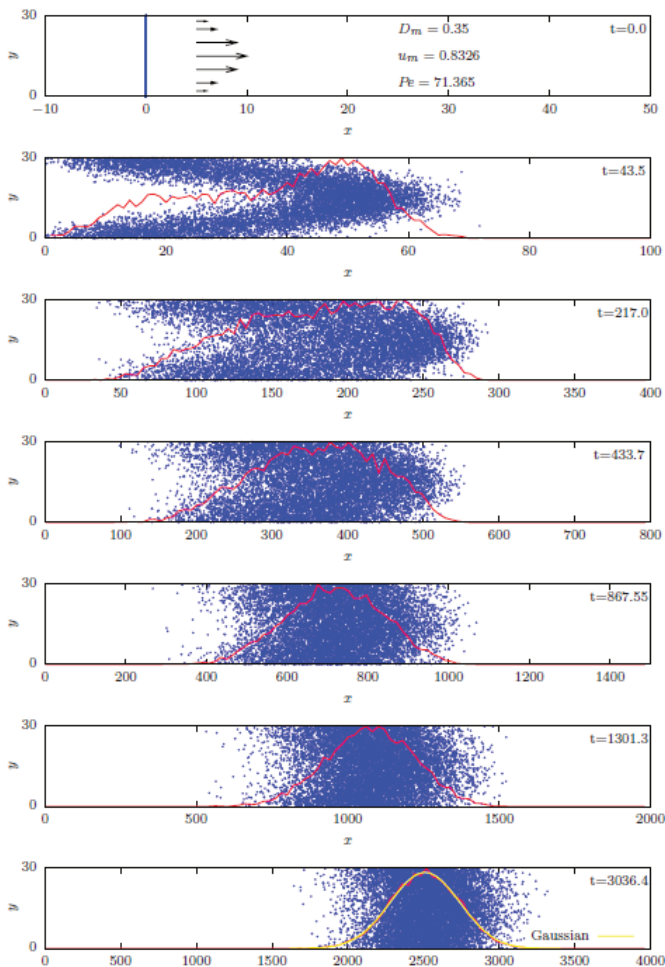
Fig. 9-2. Simulation of solute molecules transported in a capillary tube by advection and diffusion. The red line in each panel shows the mean solute concentration for each position along the tube. The yellow line in the bottom panel shows a Gaussian (normal) distribution. Source: https://www.slideshare.net/sreejithpk78/ phd–defence–3778022

In Poiseuille's classic work, which we considered in Chapter 4, he

showed that flow velocities in a capillary tube have a parabolic distribution with no flow at the tube walls and maximum velocity at the center of the tube. This velocity distribution is indicated by the black arrows on the left hand side of the top panel. When solute is first introduced at a plane in the tube (blue line, top panel) its distribution is impacted by and takes on the shape of that velocity distribution (second panel). The resulting strong concentration gradients in the direction perpendicular to the flow drive strong diffusive transport of solute. This perpendicular diffusive transport has the effect of consolidating and homogenizing the solute distribution over time (panels 3-6), until the distribution becomes a symmetrical, slowly-expanding, normal distribution centered on a plane moving with the average pore water velocity (bottom panel). This rather fascinating and utterly counter-intuitive phenomenon was discovered and described by British physicist Sir Geoffrey Taylor in 1953 [1]. Although soils bear little resemblance to Taylor's capillary tubes, his representation of hydrodynamic dispersion was soon and widely adopted to describe solute transport in soil [2]. The solute flux associated with hydrodynamic dispersion, J_{hd}, has been represented as:

$$ J_{hd} = -D_h \frac{dC}{dz} $$

<div align="right">(Eq. 9-6)</div>

where D_h is the hydrodynamic dispersion coefficient.

Combining the transport processes considered so far, we obtain an expression for the total solute flux in the soil:

$$ J_s = J_a + (J_d + J_{hd}) $$

<div align="right">(Eq. 9-7)</div>

where the fluxes due to diffusion and hydrodynamic dispersion are grouped together because they cannot be readily distinguished in practice. Inserting the flux equations for each transport mechanism, we obtain:

$$J_s = q \cdot C - D\frac{dC}{dz} - D_h\frac{dC}{dz} = q \cdot C - D_e\frac{dC}{dz}$$

where D_e is an effective diffusion-dispersion coefficient. This model for the solute transport process has been widely used in soil science, hydrology, and hydrogeology. However, it is only intended to represent the transport of solutes that do not interact with the porous media.

9.1.4 Sorption

Many solutes of interest do, of course, interact with the soil, and one of the most common interactions is the process of sorption. **Sorption** is a general term for chemical and physical processes by which solutes become attached to soil particles. Examples would include the binding of an organic pesticide like glyphosate by soil organic matter or the attachment of ammonium (NH_4^+) on a soil's cation exchange sites. Solutes that undergo sorption are transported less readily than those that do not.

In the simplest cases, we can account for the effects of sorption using a linear adsorption isotherm:

$$C_a = K_d C_l$$

where C_a is the adsorbed or absorbed chemical concentration in mass of chemical per mass of dry soil, K_d is the distribution coefficient with units of volume per unit mass, and C_l is the chemical concentration in the solution. When this linear isotherm is valid, we can compute a simple retardation factor, R, to account for the effects of sorption on the solute transport. The retardation factor is given by:

$$R = 1 + \frac{\rho_b K_d}{\Theta}$$

(Eq. 9-10)

where ρ_b is the soil bulk density.

We can then introduce this retardation factor directly into Eq. 9-4 to estimate the transport time for a solute subject to sorption:

$$t = \frac{R \cdot L \cdot \Theta}{q}$$

(Eq. 9-11)

This equation shows us that the transport time for a solute undergoing sorption is R times greater than the transport time for a solute not undergoing sorption.

9.1.5 Other Processes Impacting Solute Transport

In addition to sorption, a variety of other chemical and biological processes can influence solute transport. Chemical reactions such as oxidation/reduction, dissolution/precipitation, or association/ dissociation type reactions can strongly affect solute behavior. Biological processes such as microbial degradation or plant uptake also frequently alter the fate and transport of chemical constituents in the soil. These and diverse other processes can occur simultaneously making accurate prediction of solute transport extremely challenging. Perhaps the single greatest challenge to accurate solute transport prediction arises not from these chemical and biological factors, but rather from instances of severe heterogeneity in the water flow rate, known as preferential flow.

9.2 Preferential Flow

Preferential flow is the uneven movement of water and solutes through a relatively small portion of the soil volume at relatively high flow rates allowing these substances to reach greater depth in shorter time than would be possible in a uniform flow situation. Multiple factors can trigger preferential flow, and its effects on solute transport can be dramatic. Consider for example the preferential flow paths observed in a well-structured soil from New York (Fig. 9-3). The blue dye applied with the infiltrating water moved relatively uniformly through the topsoil, but upon reaching the subsoil, the dye was channeled almost exclusively through macropore flow paths. Macropores are ubiquitous and can be formed by burrowing animals such as earthworms or gophers, by the growth and decay of plant roots, and by soil shrinkage cracks. Macropore flow is one of the most common types of preferential flow, but is certainly not the only type.

Fig. 9-3. Preferential flow in a soil near Ithaca, New York, revealed by a blue dye applied with infiltrating water. Source: http://soilandwater.bee.cornell.edu/research/pfweb/educators/intro/macroflow.htm

In addition to macropores, other causes of preferential flow include hydrophobic soil and certain types of soil layering. A soil is considered ***hydrophobic*** when its contact angle (recall Fig. 3-10) with water is >90°. Hydrophobic conditions can arise in surface or subsurface layers where the organic matter content is high or where the soil minerals have extensive organic coatings. Soil hydrophobicity is often observed in places where the soil surface receives substantial inputs of plant residues, such as on the forest floor or even on the greens of golf courses. Water and solutes do not readily enter a hydrophobic layer, rather they tend to accumulate above the hydrophobic layer until reaching a critical point. At that point the hydrophobicity is overcome in one or more distinct

locations, and the water and solutes flow preferentially through this small, newly-wetted portion of the layer.

Figure 9-4 shows a clear example of preferential flow caused by hydrophobic soil. In this photograph of a sand dune with grass cover in the Netherlands, the wetting pattern after natural rainfall is highlighted by red dye [3]. The hydrophobic surface layer resulted in downward flow through small portions of the soil leaving much of the soil dry and increasing the total penetration depth of the wetting event. The prediction of solute transport in situations like this is further complicated by the fact that hydrophobicity can be transient. Some soils are strongly hydrophobic when dry but transition to a hydrophilic state when wetted.

Fig. 9-4. A dye staining experiment reveals preferential flow paths in a hydrophobic sand dune in the Netherlands. Reproduced from Dekker and Ritsema (2000).

Soil layering can also contribute to preferential flow, especially when a finer-textured layer has a sharp interface with a substantially coarser-textured layer. If the interface between the

layers is sloping, rather than horizontal, the flow may be "funneled" in the downslope direction (Fig. 9-5). The water does not readily pass through the coarser-textured layer because the smaller pores in the overlying finer-textured layer hold water at a more negative pressure potential than the larger pores in the coarser layer. If the interface between the layers is horizontal, then "finger" flow may develop in which water and solutes penetrate the coarser-layer only in fingers which occupy a small portion of the total volume. In either case, solutes can be transported to greater depths than they would otherwise, and reliably sampling of soil or groundwater to detect pollution becomes more difficult because of the extreme heterogeneity in the flow paths.

Fig. 9-5. Blue dye highlights the preferential flow caused by a sloping coarse-textured soil layer at a field site in New York. Source: http://soilandwater.bee.cornell.edu/Research/pfweb/educators/intro/ img_funnel.htm

9.3 Groundwater Pollution

Solute transport is one of the primary mechanisms causing groundwater pollution around the world. One common groundwater contaminant of concern is nitrate, NO_3^-, which can be readily leached from agricultural soils receiving nitrogen inputs through fertilizers or manure. In the US, observations of nitrate concentrations in groundwater have been used to develop statistical models that can estimate groundwater nitrate concentrations across the nation [4]. The US Environmental Protection Agency (EPA) has set a "maximum contaminant level" on nitrate in drinking water at 10 mg L^{-1}. Model predictions (and in many cases, actual measurements) show that this level has been exceeded for many locations in the US, particularly in portions of the Great Plains, the upper Midwest, the East Coast, and California's Central Valley (Fig. 9-6). These high nitrate concentrations are associated with relatively high nitrogen application rates, high water inputs (rain plus irrigation), and well-drained or coarse-textured soils.
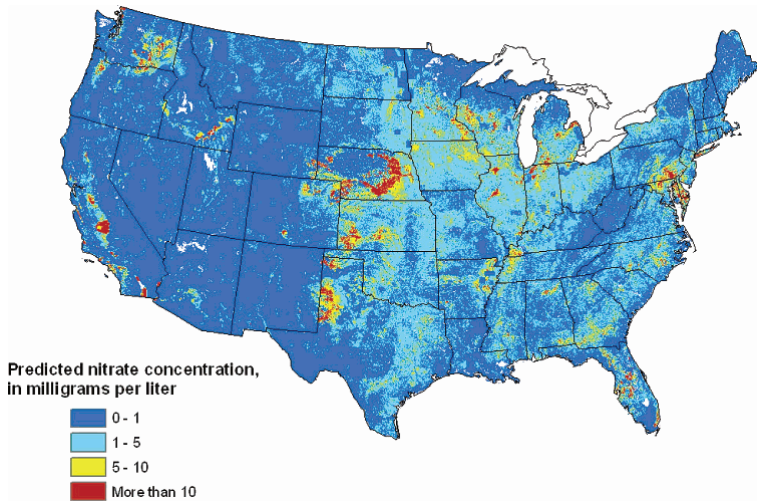
Fig. 9-6. Nitrate concentrations in shallow groundwater estimated using a nonlinear regression model considering nitrogen inputs and factors affecting nitrate transport and attenuation. Reproduced from Nolan and Hitt (2006).

Agriculture is not the only human activity which has contributed to groundwater pollution. Serious groundwater pollution has also occurred from mining, oil and gas production, industrial activities, and leaking underground storage tanks at gas stations like the ones many of us use on a regular basis. These types of contamination sometimes necessitate costly remediation efforts. Hydrocarbons such as benzene and toluene and chlorinated solvents like trichloroethylene (TCE) and tetrachloroethylene (PCE) are examples of serious groundwater contaminants arising from these types of human activities. In recent years, there has been growing concern about the process of hydraulic fracturing or "fracking" used in oil and gas production and its potential threat to groundwater resources. The US EPA has thus far not found widespread evidence of groundwater pollution due to "fracking", although some isolated cases have been reported. The available evidence indicates that accidental spillage of fracking fluid or produced water on the land

surface constitutes a greater risk to groundwater than the associated below-ground activities. This fact reinforces the importance of understanding the mechanisms of solute transport in the soil.

9.4 References

1. Taylor, G., *Dispersion of Soluble Matter in Solvent Flowing Slowly through a Tube.* Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 1953. **219**(1137): p. 186-203.
2. Elrick, D., *Dispersion and reaction in unsaturated soils applications to tracers.* Hydrological Sciences Journal, 1969. **14**(2): p. 49-60.
3. Dekker, L.W. and C.J. Ritsema, *Wetting patterns and moisture variability in water repellent Dutch soils.* Journal of Hydrology, 2000. **231–232**: p. 148-164.
4. Nolan, B.T. and K.J. Hitt, *Vulnerability of shallow groundwater and drinking-water wells to nitrate in the United States.* Environmental science & technology, 2006. **40**(24): p. 7834-7840.

10. EVAPORATION AND WIND EROSION

Thus far we have studied the spatial patterns and organization of soils across multiple spatial scales, and we have examined how soils interact with water balance processes such as precipitation, rainfall interception, raindrop impact, infiltration, runoff, water erosion, soil water redistribution, drainage, solute transport, and groundwater pollution. Now, we are ready to move toward the right hand side of Fig. 10-1 to consider the processes connected to the energy balance at the land surface.
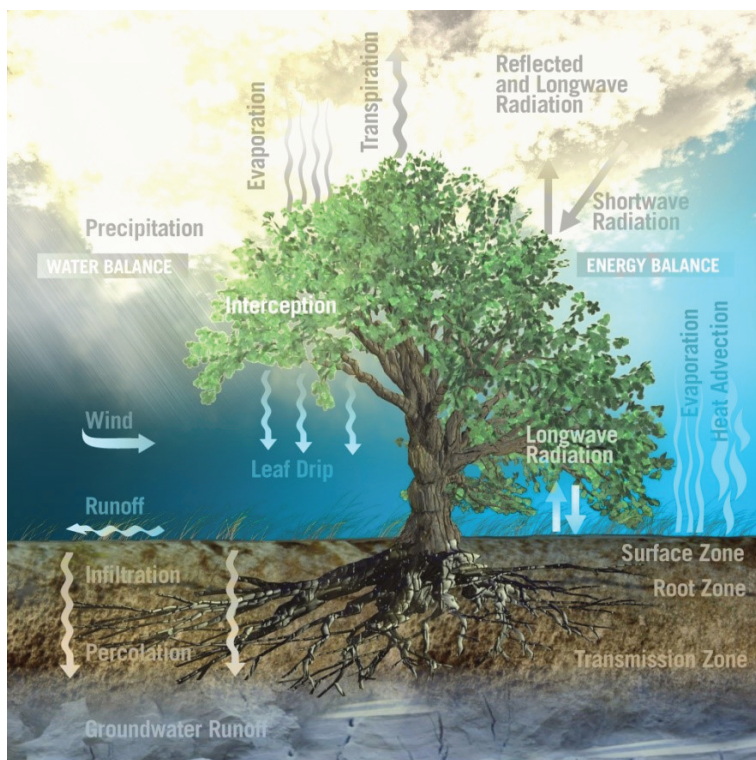
Fig. 10-1. The processes of the soil water balance (left side) and the land surface energy balance (right side). Source: European Space Agency (link).

In this chapter, we will focus on the process of evaporation. In this context, **evaporation** is the process in which water changes phase from liquid to vapor and is transported to the atmosphere from one of the following sources:

- the soil,
- the exterior surface of plants,
- plant residue, or
- surface water bodies.

Defined in this way, evaporation is distinct from the process of **transpiration**, which is the vaporization of water from the interior of plants, predominantly exiting through the stomates. The above definition also distinguishes evaporation from the composite process called **evapotranspiration**, which is simply the sum of evaporation and transpiration for a specified region of the Earth's surface.

10.1 Necessary Conditions for Evaporation

Sustained evaporation requires three things:

- a supply of energy,
- a supply of water, and
- a transport mechanism.

The substantial energy requirement for evaporation is due to one of water's peculiar properties—its unusually high latent heat of vaporization. The **latent heat of vaporization** is the energy input required to overcome the molecular forces of attraction between water molecules in liquid form. The term "latent heat" here refers to the energy that is absorbed by water during the phase change apart from any change in temperature. The value of the latent heat of vaporization is temperature dependent and at 15°C is 2.5 x 106 J kg^{-1}. Recall from Fig. 1-1 that in the SI system, J stands for joule, which is the basic unit of work or energy. Prior to adoption of the SI system, energy was often measured in calories (cal), where a calorie was the amount of energy required to raise the temperature of one gram of water by one degree Celsius at atmospheric pressure. One joule is approximately 4.18 calories. A different unit, the large calorie or kilogram calorie or food calorie (Cal), is still often used to report the energy content of foods. One food calorie is approximately 4,180 joules.

When you think of evaporation, you likely think of the energy supply for evaporation as coming from the sun. And, that is typically the primary energy source. But, it is important to recognize that the energy supply can also be drawn from the body undergoing evaporation, e.g. the soil, or from its surroundings, e.g. the air. That

is why sweat evaporating from our body helps to cool us down. That is also why evaporation can occur even during the nighttime.

In addition to a supply of energy, sustained evaporation also requires a supply of water. For soil, this means that liquid water must be transported from within the soil to the location where vaporization is occurring, which is typically at or near the soil surface. For this reason, the soil hydraulic properties which influence water flow and retention can also influence the evaporation process. The final requirement for sustained evaporation is a water vapor transport mechanism. The concentration of water vapor in the atmosphere must be less than concentration of water vapor at the evaporating surface. This concentration gradient drives diffusion of water vapor into the atmosphere. Once in the atmosphere, the vapor transport is often dominated by advection, i.e. vapor movement by bulk air flow. Thus, higher wind speeds can increase the rate of evaporation.

The above-ground supply of energy and the capacity of the atmosphere to transport vapor away from the land surface are often lumped together and described as the ***evaporative demand***. The rate of evaporation from the soil is then limited by either the evaporative demand or by the soil itself, by its ability to transport water to the location of vaporization.

10.2 Evaporation from a Water Tablehttps://open.library.okstate.edu/rainorshine/wp-admin/post.php?post=394&action=edit

In locations with a relatively shallow water table, evaporation from the soil can proceed for long periods of time and can dramatically impact the water balance and solute transport. In fact, evaporation from soils with shallow water tables is one of the major factors in the global crisis of soil salinization. When water is transported to the soil surface and evaporates, it leaves behind the salts it carried and those salts accumulate over time (Fig. 10-2). **Salinization** is the accumulation of salts in the soil to a level that negatively impacts agricultural production, ecosystem health, and economic welfare [1]. Soil salinization contributed to the downfall of ancient societies in Mesopotamia [2], and it currently affects approximately 397 Mha worldwide or 3.1% of Earth's land area [3]. Understanding the physics of evaporation from a shallow water table can help us better understand the related process of salinization.

Fig. 10-2. Soil salinization on rangeland in Colorado, USA. A layer of salt covers the soil surface and coats the wooden fence post. Public domain image. Source: https://en.wikipedia.org/wiki/Soil_salinity#/media/ File:Salinity.jpg

In the simplest approximation, we can treat evaporation from a shallow water table as a steady-state process, meaning that the rate of water movement is assumed to be constant over time and the soil is neither drying nor wetting. All the soil water that evaporates is assumed to be replenished by upward flow from the underlying groundwater table. To estimate the rate of evaporation in this case, we will apply the Buckingham-Darcy Law, Eq. 4-5.

$$q = -K(\theta)\frac{d(\Psi_p + \Psi_g)}{dz}$$

(Eq. 10-1)

To apply the Buckingham-Darcy Law in this case, we define using Campbell's hydraulic conductivity function written in terms of pressure potential:

$$\frac{K(\theta)}{K_s} = \left(\frac{\Psi_e}{\Psi_p}\right)^{2+3/b}$$

where ψ_e is the soil's air-entry potential. We define z as positive down and the water table depth as L, and we assume the pressure potential at the soil surface is $-\infty$, allowing us to integrate Eq. 10-1 to find the maximum possible steady evaporation rate, E_{max}, for a given soil and water table depth. The resulting equation is:

$$E_{max} = K_s \left[\frac{-\pi\Psi_e}{LN sin(\pi/N)}\right]^N$$

where N = 2 + 3/b. To see how we arrive at this equation and to see an example of how it can be applied, please watch this video and follow along with paper, pencil, and calculator, working the example.

10.3 Evaporation in the Absence of a Water Table

When there is no near-surface water source, like a shallow water table, then evaporation cannot be steady state. Rather, as evaporation proceeds, the soil dries out, and evaporation eventually slows down. This transient, or time-varying, evaporation is the most common type of evaporation and is often divided into three stages:

- the first constant-rate stage,
- the second falling-rate stage, and
- the third low-rate stage.

During the first stage, the evaporation is limited by the evaporative demand, and the soil is able to transmit water to the soil surface at a rate adequate to meet that demand. If the evaporative demand is constant during this stage, then the evaporation rate will be constant as well. That is the situation illustrated in the laboratory data shown in Fig. 10-3 [4]. However, in field situations the evaporative demand typically follows a daily cycle, with peak demand near midday and minimum values at night. Thus, in the field first stage evaporation is not truly constant.

The second stage of evaporation begins when the soil is no longer able to transmit water to the surface at a rate adequate to keep up with the evaporative demand. At the moment, the evaporation process switches from demand-limited to soil-limited, and the evaporation rate begins to fall (Fig. 10-3). The third stage of evaporation occurs after the soil surface has become so dry that the hydraulic conductivity is essentially zero. Since the soil can no longer transmit liquid water to the surface, the evaporation process can only proceed by sub-surface evaporation, followed by water vapor transport through the soil surface layer to the atmosphere.

During this low-rate stage, evaporation is reduced to a small fraction of the rate observed when the surface was wet.
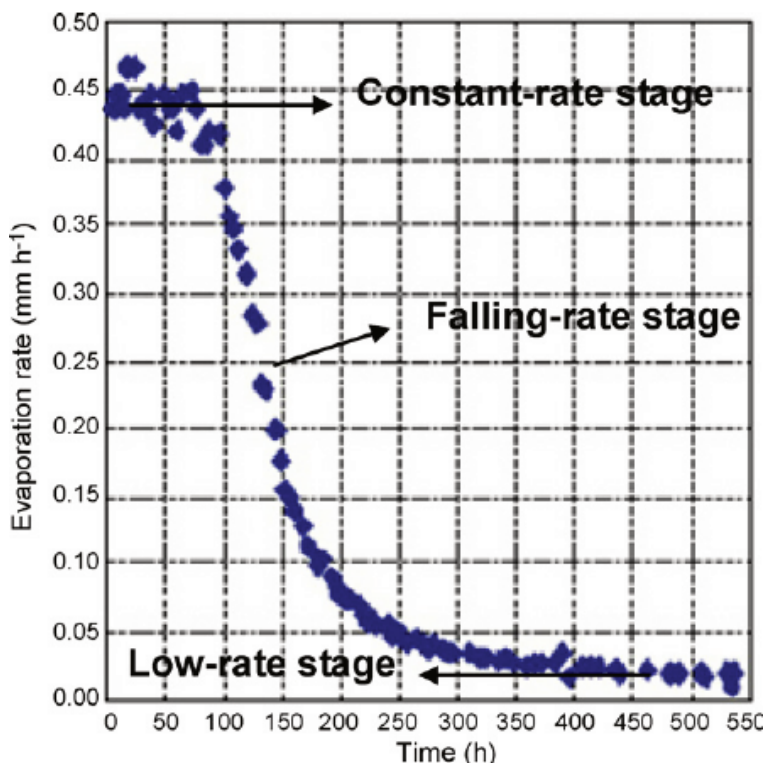

Fig. 10-3. Change of soil evaporation rate over time under constant radiation and temperature. Reproduced from Zhao et al. (2010).

The duration of each stage of evaporation depends on the soil hydraulic properties and the evaporative demand. When the evaporative demand is high, the first stage may last only a few hours and the second stage only a few days. When the evaporative demand is lower, the first stage may last for days and the second stage for weeks.

Again, it is worthwhile to emphasize that strong daily cycles of

evaporative demand often result in repeated drying and rewetting of the soil surface and in evaporation dynamics more complex than the three distinct stages just described. In a classic 1971 field experiment near Phoenix, Arizona, a loam soil was irrigated, and then detailed measurements of soil water content versus depth were recorded by soil sampling every 30 minutes for approximately 15 days [5]! A portion of this remarkable data set is shown in Fig. 10-4, and the pronounced daily wetting and drying of the soil surface is striking. Because of the strong daily cycles, the author concluded that the concept of three stages of soil evaporation appeared to have little meaning under these field conditions.
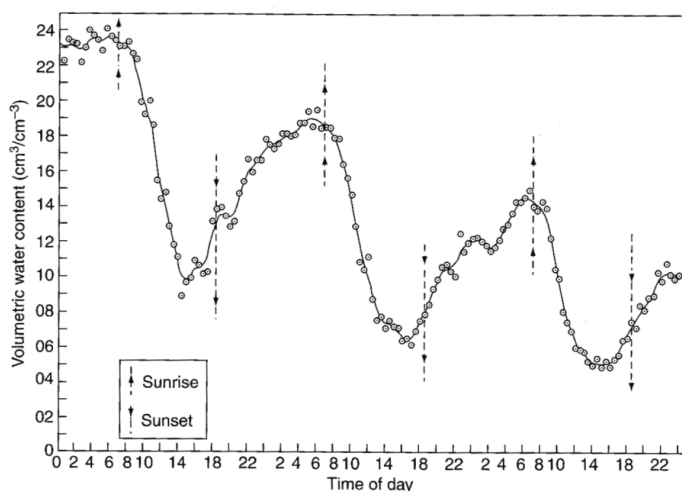
Fig. 10-4. Volumetric water content in the 0–5 mm layer of a loam soil in Arizona, USA, on days 5–7 after irrigation. The symbols are measured values and the solid line is the smoothed data. Adapted from Jackson (1973).

10.4 Reducing Evaporative Losses

Evaporation from the soil can consume a substantial portion of the available water in cropping systems, and farmers have often sought methods to reduce evaporative losses. One effective method to conserve water during the first stage of evaporation is to keep the soil covered with crop residues. Crop residues on the surface reduce the amount of solar radiation reaching the surface and can reduce the rate of water vapor transport away from the surface, both of which lower the evaporative demand. The data in Fig. 10-5 illustrate how cotton, sorghum, and wheat residues have similar beneficial effects of reducing stage 1 evaporation rate when present in equivalent thicknesses [6]. The water conserving effects of crop residue may be offset, however, because lowering stage 1 evaporation rates may lead to prolonging stage 1.
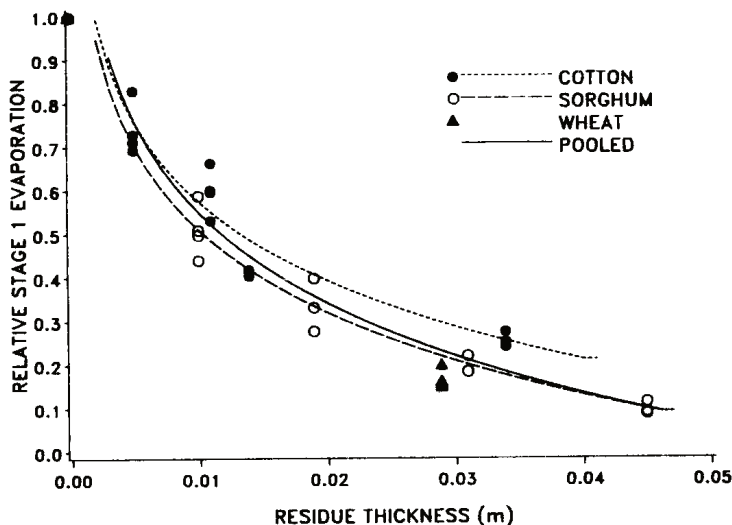
Fig. 10-5. Effect of residue thickness on first-stage evaporation. Thickness here was calculated based on the residue mass per unit area divided by the residue density. Reproduced from Steiner (1989).

In irrigated cropping systems and landscapes, evaporative losses can be reduced by decreasing the frequency of surface applied irrigation. Every time the surface is wetted, evaporation returns to first-stage levels. Applying the same irrigation amount, but in larger doses with greater time intervals between can reduce the time spent in first stage evaporation and thus reduce the cumulative evaporation. Even greater reductions in soil evaporation may be obtained by using subsurface irrigation instead of surface irrigation. A study in Texas found that up to 10% of the total seasonal water inputs might be saved due to reduced evaporation when using drip emitters at the 30-cm depth instead of surface drip emitters [7].

A variety of other methods for reducing evaporation have been tried, and some, like plastic mulches, have proven quite effective, while others have proven ill-conceived. Perhaps the most infamous method designed to reduce evaporation was the practice which

some called "dust mulch", a practice that contributed to one of the greatest natural disasters in the history of the United States.

10.5 Wind Erosion

In the early part of the 20th century, settlers dramatically expanded the area of cropland in the semi-arid western Great Plains. Encouraged by several years of favorable weather and commodity prices, these farmers plowed huge tracts of prairie to grow crops such as winter wheat. Herbicides were not available, so tillage was the primary means of weed control. Tillage was also considered by some to provide a means of reducing evaporative losses and conserving soil water. They reasoned that if a shallow surface layer of the soil was pulverized and desiccated by tillage, then the hydraulic conductivity of that layer would be so low, that soil water in the underlying layers would be protected from evaporation. This shallow, dry, pulverized layer of soil on the surface was sometimes referred to as a "dust mulch", and it was effective in reducing evaporation in some cases. But, during the severe drought of the 1930s, these aggressive tillage practices made the land vulnerable to wind erosion on a massive scale.

Fig. 10-6. A farmer and his two sons during a dust storm in Cimarron County, Oklahoma, 1936. Photo: Arthur Rothstein.

Drought and poor land management practices combined to cause the "Dust Bowl" in the US Great Plains, a time of severe wind erosion. The devastation caused by this disaster is reflected in the iconic photograph from the Oklahoma Panhandle shown in Fig. 10-6. Notice that the drifts of dust have buried large parts of the fence and the farm building, and the sky is filled with the same dust which is blowing in waves along the surface. This farm family appears to have no crops and no livestock remaining, and perhaps no hope. Due to drought, dust storms, and economic depression, millions of people packed up and moved out of the Great Plains states between

1930-1940, one of the largest migrations in US history. Although wind erosion in the US has not reached Dust Bowl levels since that time, severe dust storms have occurred in the Great Plains states as recently as 2012. Therefore, understanding and minimizing wind erosion is important still today.

As with water erosion, wind erosion includes three stages: detachment, transport, and deposition. Detachment typically occurs due to wind gusts that reach adequate velocity to separate particles from the soil surface. That is why management practices such as cover crops or standing crop residue can be effective preventive measures, since they reduce wind speed near the soil surface. Once particles are detached, they may be transported by processes such as:

- surface creep, i.e. rolling along the surface,
- saltation, i.e. bouncing along the surface, or
- suspension, i.e. being carried up and away from the surface for a substantial distance.

Deposition of windblown sediment is typically initiated by a decrease in the wind velocity, just as we saw for water-borne sediment. During the Dust Bowl barbed wire fences provided enough of a wind break to allow large drifts to form, so that in some cases livestock simply walked right over the top of the fence and wandered away. Sophisticated models have been developed to predict wind erosion processes, and one of the most widely-used models is the USDA Wind Erosion Prediction System (WEPS) [8].

10.6 References

1. Rengasamy, P., *World salinization with emphasis on Australia.* Journal of Experimental Botany, 2006. **57**(5): p. 1017-1023.

2. Jacobsen, T. and R.M. Adams, *Salt and silt in ancient Mesopotamian agriculture.* Science, 1958. **128**(3334): p. 1251-1258.

3. Setia, R., et al., *Soil salinity decreases global soil organic carbon stocks.* Science of The Total Environment, 2013. **465**: p. 267-272.

4. Zhao, S., et al., *Remote detection of bare soil moisture using a surface-temperature-based soil evaporation transfer coefficient.* International Journal of Applied Earth Observation and Geoinformation, 2010. **12**(5): p. 351-358.

5. Jackson, R.D., *Diurnal changes in soil water content during drying.* Field soil water regime, 1973(fieldsoilwaterr): p. 37-55.

6. Steiner, J., *Tillage and surface residue effects on evaporation from soils.* Soil Science Society of America Journal, 1989. **53**(3): p. 911-916.

7. Evett, S.R., T.A. Howell, and A.D. Schneider. *Energy and water balances for surface and subsurface drip irrigated corn.* Microirrigation for a Changing World: Conserving Resources/ Preserving the Environment. Proc. Fifth International Microirrigation Congress. Am. Soc. Agric. Engr., St. Joseph, MI. 1995.

8. Wagner, L.E., *A history of Wind Erosion Prediction Models in the United States Department of Agriculture: The Wind Erosion Prediction System (WEPS).* Aeolian Research, 2013. **10**: p. 9-24.

11. TRANSPIRATION AND ROOT WATER UPTAKE

In the previous chapter we focused on the process of evaporation, which is the first process we have considered in connection with the energy balance at the land surface. We noted that evaporation, as we define it, is distinct from **transpiration**, which is the vaporization of water from the interior of plants, predominantly exiting through the stomates. In this chapter we will focus on transpiration and the related process of root water uptake. Like evaporation, transpiration is a process that links together the soil water balance and the surface energy balance.

Water uptake and transport
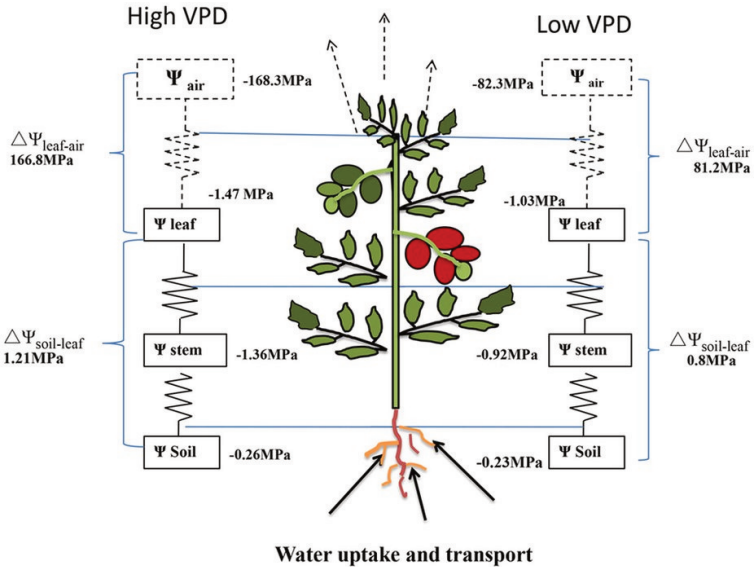
Fig. 11-1. Water potentials along the soil–plant–atmosphere continuum for greenhouse tomatoes grown under high vapor pressure deficit (left) versus low vapor pressure deficit (right) conditions. The solid and dotted lines represent the series of pathways of water flow in liquid and vapor phase, respectively. $\triangle \Psi$ represents the water potential difference between the two compartments of the soil–plant–atmospheric continuum. Reproduced from Zhang et al. (2017).

11.1 Soil-Plant-Atmosphere Continuum

We need to understand transpiration and root water uptake within the context of the ***soil-plant-atmosphere continuum***, the continuous pathway by which water moves from the soil, through plants, to the atmosphere. Water moves along this pathway in response to the typically large difference in water potential between the soil and the atmosphere. For example, in a typical greenhouse in Shaanxi, China, the air had a water potential of -168 MPa while the soil water potential was -0.26 Mpa (Fig. 11-1, left side). When the air in the greenhouse was humidified using a fogging system, the water potential of the air rose to -82 MPa (Fig. 11-1, right side), while the soil water potential was essentially unchanged. As a result the water potential difference driving transpiration was reduced by ~50% and the cumulative transpiration of the tomato plants being grown in the greenhouse was reduced by 20% [1]. This example illustrates that transpiration is, in one sense, not an active physiological function of plants but rather a primarily passive concession plants must make to the almost insatiable atmospheric demand for water.
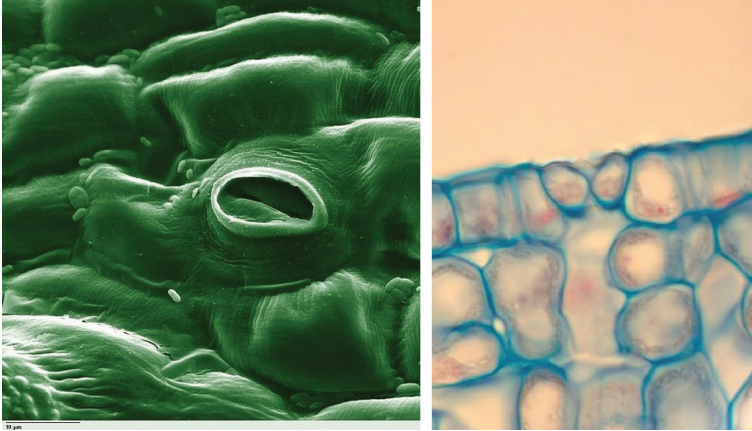
Fig. 11-2. A stomate in a tomato leaf shown using a colorized scanning electron microscope image (left), and cross-sectional view of a stomate with the guard cells which regulate the stomatal opening clearly visible (right). Both images taken from https://en.wikipedia.org/wiki/Stoma.

To understand the physics of transpiration, we need to consider the processes occurring at the stomatal level. Plant stomates are pores or openings on the exterior surfaces of leaves and other plant organs (Fig. 11-2, left side). A pair of specialized cells, called guard cells, allow the plant to open and close the stomates (Fig. 11-2, right side). The opening and closing of the stomates regulate the flow of gases such as water vapor and carbon dioxide into and out of the plant (Fig. 11-3). The plant must open the stomates to take in carbon dioxide from the atmosphere, the first step in the process of photosynthesis. But in so doing, the plant unavoidably loses water vapor to the atmosphere. The oxygen gas generated in the process of photosynthesis also exits the plant through the stomates.
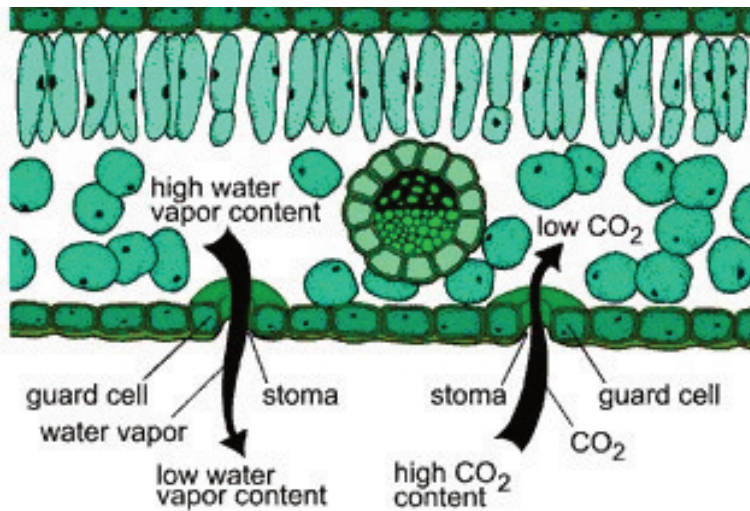
Fig. 11-3. Sketch of water vapor and carbon dioxide exchange through stomates in a plant leaf. Reproduced from http://webprojects.oit.ncsu.edu/ project/bio181de/Lab/transport/transport1.html.

11.2 Water Status of Plants

This exodus of water vapor from the plant through transpiration must be balanced by root water uptake in order for the plant to maintain a healthy water status. When the rate of transpiration exceeds the rate of root water uptake, then the water stored in the plant tissues begins to be depleted, and the plant begins to shrink or wilt. If this depletion is prolonged, the water potential inside the plant also decreases and the plant reduces the size of the stomatal openings in an effort to reduce its transpiration rate. You may have seen these processes displayed through rolled or wilted leaves of plants exposed to high evaporative demands. These are telltale signs of plant water stress.

When the evaporative demand is reduced, when the sun goes down, for example, the root water uptake rate may exceed the transpiration rate. Then, the water stores in the wilted plant are replenished, the water potential inside the plant rises, and the plant turgor is restored. The same plant recovery processes may occur, even with ongoing evaporative demand, if the soil water potential is increased due to new water inputs from rainfall or irrigation. However, under prolonged drying conditions, the plant water stress may reach a critical level of severity at which the plant cannot recover even if the evaporative demand is removed or the soil is rewetted. At this point, the plant is permanently wilted, and the soil volumetric water content corresponding to this condition is called the ***permanent wilting point***. The exact soil moisture state when permanent wilting point occurs depends on the plant species and the evaporative demand, but the soil water content at a matric potential of -1500 kPa is often used as an estimate of permanent wilting point.

11.3 Root Water Uptake

The rate of root water uptake can be limited by either the hydraulic conductivity of the rhizosphere soil, i.e. the soil immediately adjacent to the roots, or by the water potential gradient between the soil and the roots. The traditional conceptual model of root water uptake held that root water uptake lowered the water content of the rhizosphere soil, which increased the hydraulic gradient between the rhizosphere soil and the bulk soil. This increased gradient would act to increase water flow toward the rhizosphere and roots. However, the lowered water content of the rhizosphere soil also decreased the soil hydraulic conductivity (recall Fig. 4-7), which would act to decrease water flow to the roots. As long as the effect of the increased gradient was adequate to offset the effect of the decreased hydraulic conductivity, then root water uptake could proceed at a steady rate. But, inevitably, a time would come when the water potential of the rhizosphere soil reached near-equilibrium with the water potential of the roots. From that time on, the hydraulic gradient could only decrease and the hydraulic conductivity would continue to decrease. Thus, the rate of root water uptake would start to decline sharply. This is analogous, in many ways, to the transition from the constant-rate stage of evaporation to the falling-rate stage.

11.4 Transpiration and Soil Water

Since we cannot easily monitor the water status of the rhizosphere soil itself, for practical purposes, we often seek to understand and manage the relationship between transpiration and the bulk soil water. Several of the main features of that relationship are reflected in the data shown in Fig. 11-4, which is taken from a classic study on transpiration rates of corn [2]. The experiment involved corn grown in the field in 136 large containers that were individually watered to achieve a range of soil water contents for every day of the experiment. As the weather conditions, and evaporative demand, varied from day-to-day, transpiration rates for the well-watered corn varied from as low as 1.4 mm d^{-1} up to ~7 mm d^{-1}. The relative transpiration rate was defined as the ratio of the transpiration rate for the corn in a particular container to the transpiration rate of the corn in the well-watered containers on the same day. When the soil water content fell below a certain level, the relative transpiration rate decreased approximately linearly with further declines in soil water content. However, this critical soil water content value depended upon the evaporative demand, as indicated by the transpiration rate of the well-watered corn. When the well-watered transpiration rate was 6.4 mm d^{-1}, the relative transpiration rate declined once soil water contents dropped below 0.34 cm^3 cm^{-3} in this Colo silty clay loam. Under conditions of lower evaporative demand when the well-watered transpiration rate was 2.0 mm d^{-1}, the relative transpiration rate did not decline until soil water content dropped below 0.25 cm^3 cm^{-3}. These data show how both evaporative demand and soil water status interactively influence the process of transpiration.
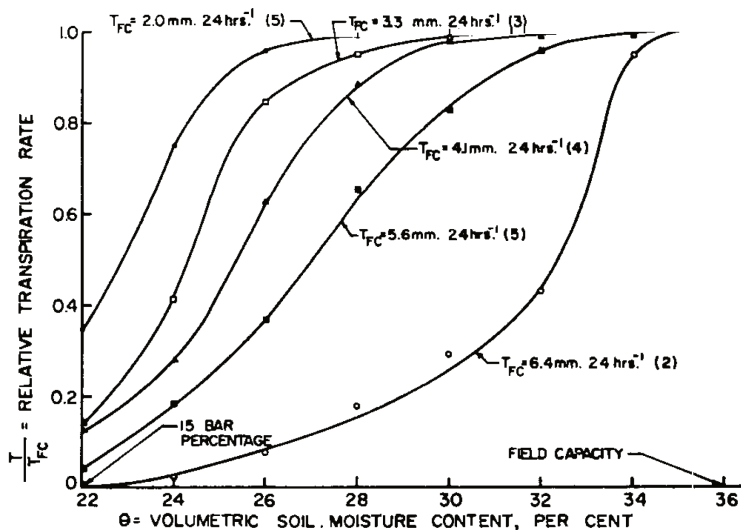
Fig. 11-4. Transpiration rate (T) of corn plants under water stress relative to that of well-watered corn plants (T_{FC}) as a function of soil volumetric water content (θ). Curves represent different levels of T_{FC} increasing from 2 mm d^{-1} (leftmost curve) to 6.4 mm d^{-1} (rightmost curve). 15 bar = –1500 kPa. Reproduced from Denmead and Shaw (1962).

11.5 Soil Water Availability Indicators

The specific soil water content at which the relative transpiration rate begins to decline depends, not only on the evaporative demand (and on the plant species), but also on the soil water retention characteristics. Therefore, when considering transpiration and plant water use, we often adjust or normalize the soil water content values to account for some of the soil specific differences. One common adjustment is to compute the **plant available water** (PAW), which is the equivalent depth of water which is available for plant uptake within a layer of soil of specified thickness. The value of PAW for a soil layer is:

$$PAW = (\Theta - \Theta_{wp}) \cdot L$$

where θ is the current value of soil volumetric water content, θ_{wp} is the soil volumetric water content at the permanent wilting point, and L is the thickness of the soil layer. Plant available water has length units, e.g. mm, the same as precipitation, which makes PAW values relatively easy to understand. Because PAW accounts for the soil-specific permanent wilting point, it provides a better indicator of water status across different soil types than does volumetric water content.

The capacity of a specific soil to store plant available water, i.e. the **available water capacity** (AWC), is often approximated by:

$$AWC = (\Theta_{fc} - \Theta_{wp}) \cdot L$$

where θ_{fc} is the volumetric water content corresponding to "field capacity". Despite the previously discussed shortcomings of the field capacity concept, AWC has proven to be a practical indicator of differences between soils in their capacity to store water for plant

uptake. We can combine the prior two equations to determine the *fraction of available water capacity* (FAW) which is filled for at any particular value of soil water content:

$$FAW = \frac{\Theta - \Theta_{wp}}{\Theta_{fc} - \Theta_{wp}}$$

(Eq. 11-3)

Although there is substantial variability in the value of FAW below which plant water stress occurs, we can often use FAW = 0.5 as an approximate threshold for plant water stress. Effective procedures for modeling plant water use and adjusting water stress thresholds for specific situations are available [3].

If we have soil volumetric water content measurements, and estimates of the soil water retention properties, then we can use the FAW concept to estimate or monitor the degree of plant water stress. For example, FAW has proven to be a good indicator of the probability of large wildfires during the growing season in Oklahoma [4]. These large growing season fires only occur if the vegetation is under water stress severe enough to reduce the vegetation water content to levels low enough that the vegetation will readily burn. The vast majority of large growing-season fires in Oklahoma occur when FAW drops below 0.2, indicating severe water stress.
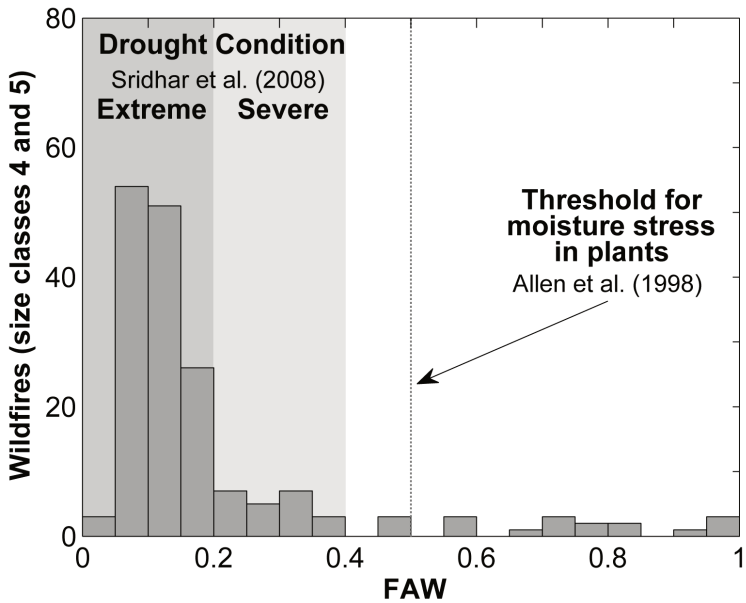
Fig. 11-5. *Frequency distribution of fraction of available water capacity (FAW) for growing–season wildfires >121 ha in size in Oklahoma from 2000–2012. Most fires (159 of 174 fires for which soil moisture data were available) occurred at FAW < 0.5, with some occurring under severe drought (22 fires) and most occurring under extreme drought (134 fires). Reproduced from Krueger et al. (2015).*

11.6 Water Use Efficiency

As we saw in Fig. 11-3, the process of transpiration is inextricably linked to the process of carbon assimilation through photosynthesis. Thus, transpiration is a critical link between three of the Earth's most important cycles or balances, the soil water balance, the surface energy balance, and the atmospheric carbon balance. A great deal of research attention and even public interest has been focused on one aspect of that linkage—the water use efficiency of plants. Water use efficiency is an important concept that can be easily misunderstood and overemphasized.

Plant **water use efficiency** (WUE) is the ratio of carbon or biomass accumulation to water use or supply during a given time span. One reason that WUE is easily misunderstood is the wide variety of ways in which it can be defined. The accumulation can be expressed in terms of:

- carbon dioxide assimilation,
- above-ground biomass, or
- harvested yield.

The water consumed can be expressed in terms of:

- transpiration,
- evapotranspiration, or
- total water supply.

And, the time period of interest can be:

- instantaneous,
- seasonal, or
- annual.

The factors controlling water use efficiency in plants are relatively

well-understood, and there appears to be little scope for altering water use efficiency at a fundamental level [5]. Stomates are not selective. When they open to let carbon dioxide in, they necessarily let water vapor out (Fig. 11-3). The ratio of carbon in to water out depends primarily on the size of the respective gradients in gas concentrations. The carbon dioxide flow into the leaf is driven by the difference between carbon dioxide concentration in the atmosphere and that inside the leaf. The flow of water vapor out of the leaf is driven by the difference between the water vapor concentration inside the leaf and that in the atmosphere. Thus, the WUE expressed as the ratio of carbon assimilated (A) to transpiration (T) for a given instant (i) at the leaf-level (l) is approximately:

$$WUE(A, T, i)_l = 1.6c \frac{Pa}{(e_l^* - e)}$$

where c is a constant that depends on the ratio of the internal and external carbon dioxide concentrations, P_a is the partial pressure of carbon dioxide in the atmosphere, e^*_l is the saturation vapor pressure for water vapor at the temperature of the leaf, and e is the vapor pressure of water in the atmosphere [5]. An interactive graph to illustrate the relationships between e^*_l and e as affected by temperature and relative humidity is available here.

The primary possibilities for improving water use efficiency at the leaf level are indicated in Eq. 11-4. The constant c has a value of approximately 0.3-0.4 for C_3 plants and 0.7 for C_4 plants. This means that crop water use efficiency can be increased by using more C_4 crops and fewer C_3 crops. C_3 plants account for about 85% of all plant species, including major crops like wheat, rice, and soybean. These plants have relatively low water use efficiency in part because in these plants the carbon-fixing enzyme rubisco is exposed to relatively low internal carbon dioxide concentrations, thus its carbon-fixation is relatively inefficient. C_4 plants account for only about 3% of plant species, including corn, sorghum, and

sugarcane. These plants have a special cells that ensure rubisco is exposed to higher internal concentrations of carbon dioxide, thereby enhancing carbon fixation efficiency. As a result, C_4 plants have substantially higher water use efficiency than C_3 plants. Plants, like pineapple, that use the crassulacean acid metabolism (CAM) photosynthetic pathway achieve high water use efficiency by opening their stomates at night. A helpful comparison of the C_3, C_4, and CAM pathways is available here.
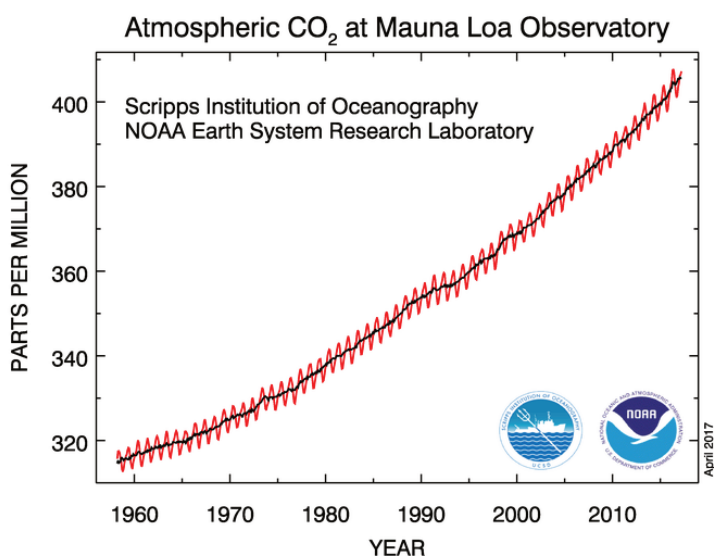
Fig. 11-6. Monthly mean atmospheric carbon dioxide at Mauna Loa Observatory, Hawaii. The carbon dioxide data (red curve), measured as the mole fraction in dry air, on Mauna Loa constitute the longest record of direct measurements of CO_2 in the atmosphere. They were started by C. David Keeling of the Scripps Institution of Oceanography in March of 1958. For the latest data click here.

A second factor impacting water use efficiency is the concentration or partial pressure (P_a, Eq. 11-4) of carbon dioxide in the atmosphere.

As the carbon dioxide concentration in the atmosphere increases, the water use efficiency of plants increases because the increased external concentration drives a greater rate of diffusion of carbon dioxide into the leaves when the stomates are open. Due mainly to fossil fuel combustion, the carbon dioxide concentration of Earth's atmosphere is steadily rising (Fig. 11-6), and, if no other factors were changing, this trend would cause increased plant water use efficiency. Because global climate change is multi-faceted, other concurrent changes such as increasing air temperatures could potentially act to reduce or prevent the expected increase in plant water use efficiency. However, examination of data from forests in the Northern Hemisphere has in fact shown a substantial increase in water use efficiency over the past two decades [6]. Available evidence suggests that trees may benefit more than other plant types [7], and that C_3 crops may benefit more than C_4 crops from the rising carbon dioxide in the atmosphere (see Table 11-1, [8]).

Table 11-1. Relative increase in water use efficiency (WUE) of field-grown crops with CO_2 enrichment above the ambient CO_2 concentration. Crops were well-watered unless noted. Adapted from Polley (2002)

Species	Measurement type	% increase in WUE	% increase in CO_2
Soybean	Gas exchange	102	142
Cotton	Total biomass		
wet		28–39	49
dry		19–37	49
Rice	Gas exchange	13–53	100
Rice	Gas exchange		
wet		34–53	100
dry		125	100
Wheat	Grain mass		
wet		40	143
dry		46	143
Wheat	Grain mass		
wet		17	49
dry		32	49
Sorghum	Total biomass	34	100

The third possibility for improving water-use efficiency is reflected,

not only in Eq. 11-4, but also in Fig. 11-1. When plant stomates are opened in times or locations where the atmospheric vapor pressure deficit, represented by $(e*_l − e_l)$ in Eq. 11-4, is relatively low, then the gradient driving diffusion of water vapor out of the leaves will be relatively small, and the WUE will be relatively high. In the context of agriculture, this would suggest that if our goal is to increase WUE, then we should consider focusing future crop production in more humid climates. Another option could be to shift crop growing seasons earlier such that more crop growth occurs during the cooler parts of the year when the vapor pressure deficit is lower. Of course, such shifts in the growing season may increase WUE but also increase risk of freeze damage or slow crop growth due to low temperatures.

Major research investments have been made around the world to study the intrinsic water use efficiency of important crops and to seek ways to increase water use efficiency through plant breeding or genetic engineering. However, the logical justification for this work may not be altogether solid [9]. There is not clear evidence that increasing the intrinsic water use efficiency of crop varieties will result in increased crop production or more drought-resistant crops or reduced water use in agriculture [10]. A stronger case can be made that research efforts should be focused on improving crop yields by increasing crop water use, i.e. transpiration, not water use efficiency per se. Crop varieties that are able to sustain higher rates of transpiration through effective use of soil water, especially during critical reproductive periods, will be more productive than other varieties in water-limited environments. This is a consequence of the strong, linear relationship of plant biomass accumulation to transpiration.

11.7 Problem Set

1. A young corn plant in a growth chamber produced 31 g of above-ground biomass with a cumulative transpiration of 4500 g of water.

   a. What was its water use efficiency (WUE) in g kg^{-1}?
   b. For the same growing conditions, would you expect higher or lower WUE for a native C_3 plant? Explain why in one sentence.
   c. If the vapor pressure deficit during the growth period were increased, would the WUE be higher or lower? Explain why in one sentence.

2. Imagine you are responsible for scheduling irrigation for a corn crop. Your management goal it to keep the corn well-watered so that the evapotranspiration from the field is equal to the reference (or potential) evapotranspiration rate. Create a spreadsheet to perform a simple daily soil water balance for the field to **determine the number of irrigation events and total irrigation amount** required in the 120-d growing season assuming:

   ◦ The reference evapotranspiration rate starts at 2.0 mm d^{-1} on day 1 of the season and increases by 0.1 mm d^{-1} each day until day 60, after which it decreases by 0.1 mm d^{-1} each day.
   ◦ Rainfall of 12 mm occurs early in the morning on day 7, day 14, and every 7th day after that for the duration of the season, and all the rain infiltrates into the soil.
   ◦ The root zone is 1-m deep silt loam with a "field capacity" of 0.23 cm^3 cm^{-3} and permanent wilting point of 0.09 cm^3 cm^{-3} and is at field capacity on day 0.
   ◦ Each irrigation event occurs on the last possible day to

maintain FAW > 0.5. The water added in each irrigation event is the amount required to restore the profile to field capacity based on the soil water storage at the end of the previous day.

3. Repeat problem 2, but this time for a loamy sand with a "field capacity" of 0.09 cm^3 cm^{-3} and permanent wilting point of 0.02 cm^3 cm^{-3}. Report the number of irrigation events and total irrigation amount required.

11.8 References

1. Zhang, D., et al., *Vapour pressure deficit control in relation to water transport and water productivity in greenhouse tomato production during summer.* Scientific Reports, 2017. **7**: p. 43461.

2. Denmead, O.T. and R.H. Shaw, *Availability of Soil Water to Plants as Affected by Soil Moisture Content and Meteorological Conditions.* Agronomy Journal, 1962. **54**(5): p. 385-390.

3. Allen, R.G., et al., *Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, in FAO, Rome.* 1998. p. D05109.

4. Krueger, E.S., et al., *Soil Moisture Affects Growing-Season Wildfire Size in the Southern Great Plains.* Soil Science Society of America Journal, 2015. **79**(6): p. 1567-1576.

5. Sinclair, T.R., C.B. Tanner, and J.M. Bennet, *Water-Use Efficiency in Crop Production.* BioScience 1984. **34**: p. 36-40.

6. Keenan, T.F., et al., *Increase in forest water-use efficiency as atmospheric carbon dioxide concentrations rise.* Nature, 2013. **499**(7458): p. 324-327.

7. Ainsworth, E.A. and S.P. Long, *What have we learned from 15 years of free-air CO2 enrichment (FACE)? A meta-analytic review of the responses of photosynthesis, canopy properties and plant production to rising CO2.* New Phytologist, 2005. **165**(2): p. 351-372.

8. Polley, H.W., *Implications of Atmospheric and Climatic Change for Crop Yield and Water Use Efficiency* Crop Science, 2002. **42**(1): p. 131-140.

9. Tanner, C.B. and T.R. Sinclair, *Efficient Water Use in Crop Production: Research or Re-Search?, in Limitations to Efficient Water Use in Crop Production*, H.M. Taylor, W.R. Jordan, and T.R. Sinclair, Editors. 1983, American Society of Agronomy, Inc.; Crop Science Society of America, Inc.; Soil Science Society of America, Inc.: Madison, WI. p. 1-27.

10. Blum, A., *Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress.* Field Crops Research, 2009. **112**(2–3): p. 119-123.

12. SURFACE ENERGY BALANCE AND EVAPOTRANSPIRATION

In the previous two chapters we focused on evaporation and transpiration, two processes which are major energy sinks at the land surface. In this chapter, we will directly examine the energy balance for the land surface, represented on the right side of Fig. 12-1. We will consider the nature and fate of radiation received at the land surface, the exchange of sensible heat between the surface, the atmosphere, and the soil, and the consumption of energy (latent heat) during evapotranspiration. The surface energy balance is essentially a way of accounting for heat transfer between the land surface and its surroundings.
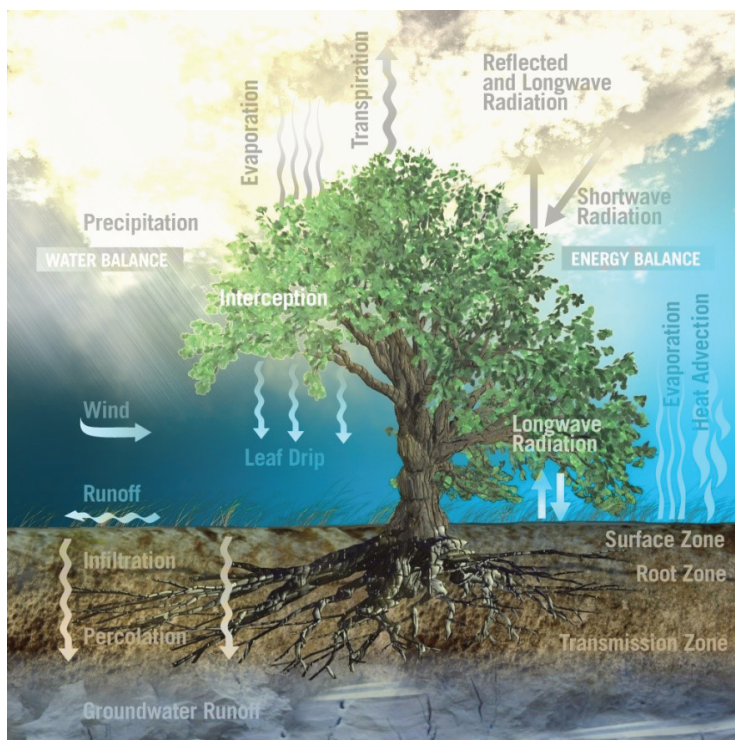
Fig. 12-1. The processes of the soil water balance (left side) and the land surface energy balance (right side). Source: European Space Agency (link).

Heat transfer is simply energy transfer that depends on the temperatures of the objects or systems involved. ***Temperature*** is a measure of the average kinetic energy of the random microscopic movements of the molecules or particles of a substance. Temperature is also the physical quantity which determines the direction of heat transfer between two substances in thermal contact with each other. Heat is transferred from regions of higher temperature to regions of lower temperature. Viewed in this way, temperature is somewhat analogous to water potential, the physical quantity which determines the direction of water flow.

12.1 Modes of Heat Transfer

To understand the surface energy balance, we need to understand the three fundamental modes of heat transfer:

- radiation,
- conduction, and
- convection.

Radiation is the emission or transfer of energy in the form of electromagnetic waves or particles. **Conduction** is heat transfer due to the random thermal motion of the molecules and particles within a system. **Convection** is heat transfer due to bulk fluid motion. Each of these modes plays an important role in the energy balance of the land surface. In the surface energy balance, the energy flux via radiation is often the largest single term, so we will consider this mode of heat transfer first.

12.2 Radiation Basics

The sun, the land, the oceans, and even the atmosphere constantly emit radiation in the form of electromagnetic waves. This radiation can be classified according to its position on the electromagnetic spectrum (Fig. 12-2). The wavelength of maximum radiation intensity (λ_{max}) emitted by an object is inversely related to the object's temperature. For a perfect emitter, or *blackbody*, Wien's displacement law describes the relationship between λ_{max} and object's absolute surface temperature, T (K):

$$\lambda_{max} = \frac{b}{T}$$

(Eq. 12-1)

where b = 2,900 μm K.
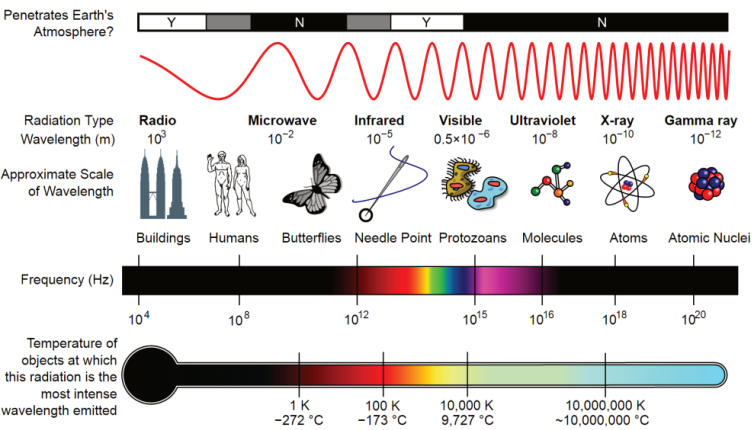
Fig. 12-2. A diagram of the electromagnetic spectrum, showing various properties across the range of frequencies and wavelengths. By Inductiveload, NASA – self-made, information by NASA. Based off of File:EM_Spectrum3-new.jpg by NASA. The butterfly icon is from the P icon set, P biology.svg. The humans are from the Pioneer plaque, Human.svg. The buildings are the Petronas towers and the Empire State Buildings, both from Skyscrapercompare.svg. CC BY-SA 3.0, (link)

The sun, with a surface temperature of approximately 5,780 K, emits radiation with a peak intensity at wavelength of approximately 0.5 μm, which is in the green range of the visible portion of the electromagnetic spectrum. The Earth's surface, which has an average temperature of approximately 287 K, emits radiation with a peak intensity of around 10 μm, which is in the infrared portion of the spectrum. Thus, in the context of the surface energy balance, the incoming solar radiation is called **shortwave radiation**, while the radiation emitted by the land surface and atmosphere is called **longwave radiation**.

The intensity of radiation emitted by an object depends on the object's temperature even more strongly than does the wavelength. The equation describing the relationship between an object's temperature and the intensity of its radiation is called the Stefan-Boltzmann Law which can be written as:

$$J_t = \varepsilon \sigma T^4$$

(Eq. 12-2)

where J_t is total radiant flux from the emitting surface (W m^{-2}), ε is the emissivity of the surface, σ is the Stefan-Boltzmann constant (5.67 x 10^{-8} W m^{-2} K^{-4}), and T is the absolute temperature of the surface (K).

The **emissivity** is a measure, ranging from 0 to 1, of the effectiveness of a surface in emitting radiation. A perfect emitter has an emissivity of 1. The sun, along with snow and ice covered surfaces are nearly perfect emitters and have emissivity values of approximately 0.99. Water has an emissivity between 0.98 and 0.99, while vegetated surfaces have emissivity values typically ranging from 0.95 to 0.98 for infrared radiation in the 8-13 mm wavelengths [1]. The emissivity of soil is positively related to soil water content [2] and varies between 0.86 and 0.96 [3]. Thus, the composite emissivity of the Earth's land surface is highest at north and south poles and lowest in the mid-latitude deserts (Fig. 12-3) [4].
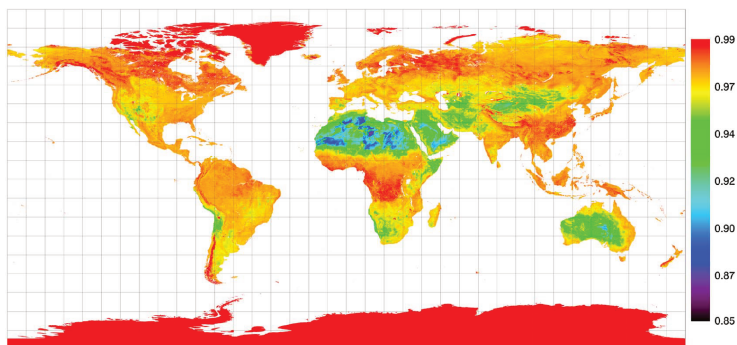
Fig. 12-3. Estimated global land surface emissivity on day of year 241 in 2008. Reproduced from Cheng and Liang (2014).

It is important to recognize that just over half of the solar radiation arriving at the top of the Earth's atmosphere actually arrives at the land surface. On average, about 25% of the incoming solar radiation is reflected back to space by clouds (Fig. 12-4). An additional 5% is reflected by air molecules, dust, or aerosols in the atmosphere. Approximately 14% of the incoming solar radiation is absorbed by air molecules, dust, or aerosols, and 3% is absorbed by clouds. Only about 31% of the solar radiation travels directly to the Earth's surface, and an additional 22% reaches the surface as *diffuse radiation* after being scattered or reflected in the atmosphere.

12.3 Net Radiation

When the direct and diffuse shortwave (solar) radiation reach the land surface, some of this radiation is reflected by the surface. The fraction of the incoming shortwave radiation that is reflected by the surface is called the ***albedo***. In general, dark surfaces have low albedo while light colored surfaces have high albedo. Fresh snow can have an albedo as high as 0.9, reflecting large portions of the incoming shortwave radiation. In contrast, the albedo of water is typically quite low, < 0.1, although it can be much higher when the sun angle is low. Soils and vegetation typically have intermediate values of albedo, ranging from 0.1 to 0.4. Moist soil surfaces are typically darker and have lower albedo values than do dry soil surfaces. Soil color is also affected by soil organic matter, with higher levels of soil organic matter resulting in darker soil and lower albedo.
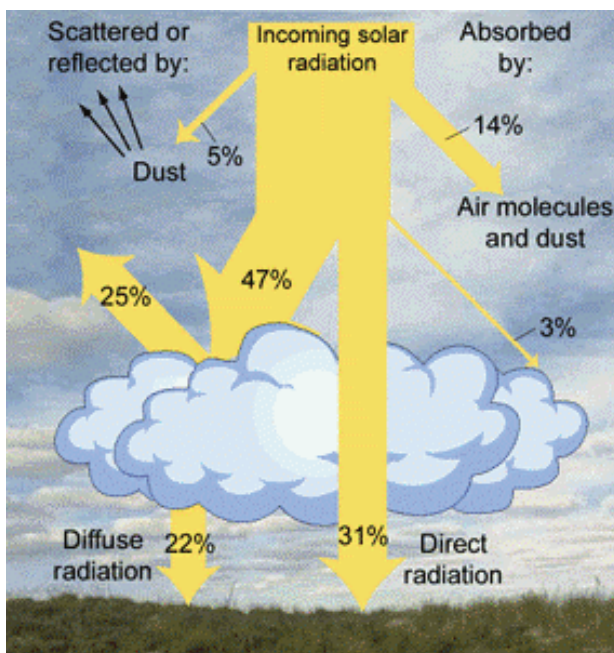
Fig. 12-4. Partitioning of solar radiation by the atmosphere.
Source: (link)

In addition to the shortwave radiation, we need to account for the longwave radiation at the land surface. We have seen previously that the land surface emits longwave radiation to the atmosphere. But, we should not overlook the fact that the atmosphere emits longwave radiation both upward into space and downward toward the Earth's surface. The magnitude of this downward longwave radiation depends on the temperature and emissivity of the atmosphere, according to Eq. 12-2. The atmospheric emissivity can range from 0.5 to nearly 1 [5]. The presence of clouds significantly increases the emissivity of the atmosphere, which is why, as you have perhaps noticed, cloudy winter nights are often warmer than clear winter nights. The amount of water vapor in the atmosphere

also has a strong effect on the atmospheric emissivity, with higher emissivity values for more humid conditions [6].

Typically the largest term in the surface energy balance is the net radiation. **Net *radiation*** is simply the sum of all incoming and outgoing radiation fluxes at the land surface. Mathematically, net radiation, R_n, is given by:

$$R_n = (1 - \alpha)R_s + R_{li} - R_{lo}$$

(Eq. 12-3)

where α is the albedo, R_s is the incoming solar radiation (direct and diffuse), R_{li} is the incoming longwave, and R_{lo} is the outgoing longwave.

12.4 Surface Energy Balance

Having defined net radiation, we can now write the surface energy balance equation:

$$R_n = LE + H + G$$

where LE is the latent heat flux, H is the sensible heat flux, and G is the soil heat flux. The fluxes are commonly expressed in $W\,m^{-2}$.

By convention, net radiation is defined as positive towards the surface, while all the other terms are defined as positive away from the surface. During the daytime, net radiation is typically towards the surface, and the other fluxes are away from the surface (Fig. 12-5). In the nighttime, the fluxes are in the opposite directions. The latent heat flux is typically near 0 during the night, although negative values, indicating condensation, are possible.

The **latent heat flux** (LE) is the energy that is absorbed by water at the Earth's surface during evaporation or transpiration apart from any change in temperature. It is a flux or transfer of energy because we assume the resulting water vapor is transported away from the surface by diffusion and advection in the atmosphere. The latent heat flux is equal to the latent heat of vaporization for water multiplied by the evapotranspiration (ET) rate.
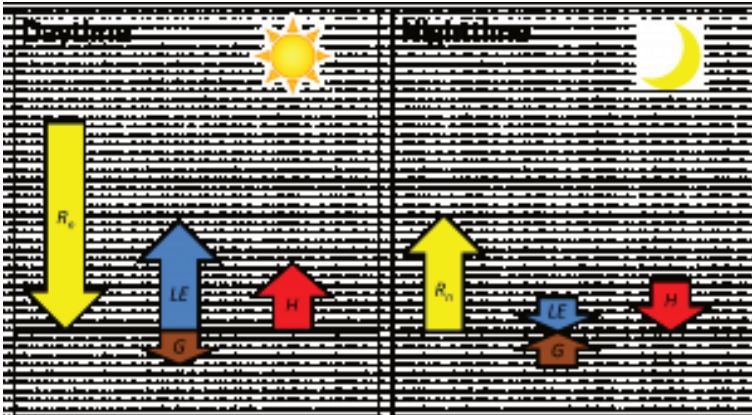
Fig. 12-5. Diagram showing hypothetical daytime and nighttime directions for each of the fluxes in the energy balance for a moist land surface in the summer.

The **sensible heat flux** (H) is the heat transfer between the surface and the atmosphere by conduction and convection. During the daytime, the land surface is often warmer than the atmosphere, so the surface heats the air. As the air comes into direct contact with the land surface, energy is transferred from the surface to the air by conduction. Then, wind moves that air away from the surface, transferring energy to the atmosphere by convection. At night, the surface is often cooler than the atmosphere because of radiative cooling, so the processes just described are reversed, and sensible heat flux is toward the surface.

The **soil heat flux** (G) is the heat transfer between the surface and the underlying soil, predominantly by conduction. Radiation is of negligible importance for subsurface heat transfer. Convective heat transfer by flowing water can be important during times and places of relatively high water flow rates, but these are not the norm. We will examine soil heat flux in more detail in the next chapter, which focuses on soil temperature.

Measurements of the components of surface energy balance

equation have been widely used by researchers to better understand the transfers of water and energy in croplands, grasslands, and forests. For example, energy balance measurements based on a Bowen ratio technique were used to determine how crop residues affected heat transfers, and ultimately soil moisture and temperature, in an Iowa corn field during the fall and spring (Fig. 12-6) [7]. During the time period shown in Fig. 12-6, the net radiation ranged from -50 W m^{-2} at night to +300 W m^{-2} around noon. The sensible heat flux also peaked around noon but at a value of +200 W m^{-2}. The crop residue suppressed evaporation and buffered the soil surface temperature, so the latent heat flux and soil heat flux were smaller components of the energy balance, not exceeding +100 W m^{-2} in this case.
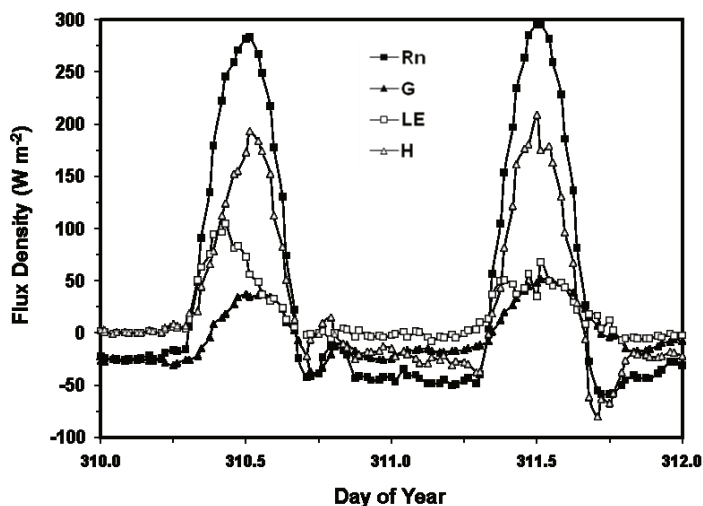
Fig. 12-6. Energy balance components measured above a corn residue covered soil surface in 1994 at a site near Ames, Iowa. Net radiation (Rn) is positive toward the surface. The other terms are positive away from the soil surface. Adapted from Sauer et al. (1998).

12.5 Reference Evapotranspiration

One of the most widespread uses of the surface energy balance concept is for determining evapotranspiration. The actual evapotranspiration rate for a given time for a particular location on the land surface can be substantially influenced by site-specific management practices. Thus, often it is useful to estimate the potential or **reference evapotranspiration** (ET_0) which is the evapotranspiration rate expected for a standard well-watered surface under the prevailing atmospheric conditions. The daily or monthly value of ET_0 can be determined for relatively large regions, and those values can be adjusted to estimate actual ET for a specific location if adequate information is available [8]. Methods for estimating ET_0 or ET commonly are based on the surface energy balance.

12.5.1   Hargreaves Method

One of the simplest approaches for estimating ET_0, the Hargreaves method, was developed to facilitate irrigation management and water planning. The initial work began in California, and the importance of having a simple approach with low data requirements was solidified when the work expanded to Haiti [9]. Building on earlier work in which ET_0 was calculated as a linear function of incoming solar radiation (R_s), Hargreaves found that the product of R_s and air temperature (T) explained much of the variance in measured ET_0. The problem was that measurements of R_s are often unavailable. Therefore, Hargreaves and colleagues developed a way to estimate Rs from knowledge of extraterrestrial radiation (R_a, solar

radiation at the top of Earth's atmosphere for a specific place and time) and air temperature [10].

Since R_a can be accurately estimated based on site latitude and day of year and air temperature measurements are widely available, the Hargreaves method can be used virtually anywhere. The Hargreaves equation is:

$$ET_0 = 0.0023 R_a (\bar{T} + 17.8)(T_{max} - T_{min})^{0.5}$$

(Eq. 12-5)

where R_a is the cumulative extraterrestrial radiation for the calculation period converted to units of mm of water, $\bar{T}$ is the mean temperature (°C) for the calculation period, T_{max} is mean daily maximum temperature for the calculation period, and T_{min} is mean daily minimum temperature for the calculation period. This equation is designed for use at 5-day or larger calculation periods and is frequently used with monthly calculation periods. The Hargreaves method is a good choice for ET_0 estimation when the only meteorological data available are air temperatures. For example, the Hargreaves method was used to estimate a 45-yr time series of ET_0 for a large watershed stretching across southwest Oklahoma and part of west Texas, USA (Fig. 12-7) [11]. Compiling long time series like this would be more difficult, or in some cases impossible, for more data intensive ET_0 estimation methods.

12.5.2 Penman-Monteith Method
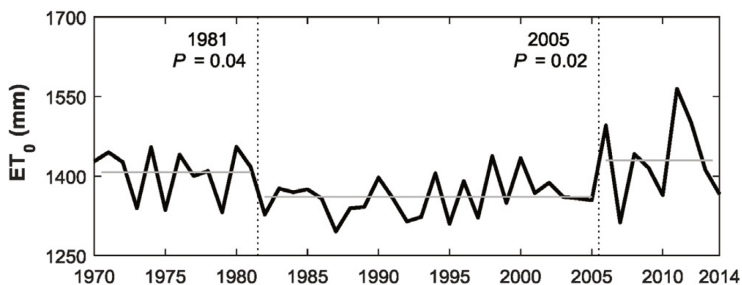
Fig. 12-7. *Annual reference evapotranspiration (ET$_0$) totals from 1970–2014 for the North Fork of the Red River watershed which lies in southwest Oklahoma and the Texas panhandle. Monthly ETO values were calculated using the Hargreaves method and summed to get the annual totals. Adapted from Krueger et al. (2017).*

A more precise approach was for estimating ET was proposed by Penman, who in 1948 combined the equations for the surface energy balance and for turbulent transport of water vapor away from a saturated evaporating surface [12]. In 1965, Monteith expanded the Penman approach to surfaces where the vapor pressure was less than the saturated vapor pressure (i.e. <100% relative humidity) [13]. The Penman-Monteith equation is:

$$ET = \frac{1}{L} \frac{\Delta(R_n - G) + \rho_a c_p \frac{e_s - e_a}{r_a}}{\Delta + \gamma \left(1 + \frac{r_s}{r_a}\right)}$$

(Eq. 12-6)

where L is the latent heat of vaporization for water (J kg^{-1}), R_n is the net radiation (W m^{-2}), G is the soil heat flux (W m^{-2}), $e_s - e_a$ is the vapor pressure deficit of the air (kPa), r_a is the density of the air (kg m^{-3}), c_p is the specific heat capacity of the air at constant pressure (J kg^{-1} °C^{-1}), Δ is the slope of the relationship between saturated vapor

pressure and temperature (kPa °C^{-1}), γ is called the psychrometric constant (kPa °C^{-1}), r_s is the surface resistance to vapor transport (s m^{-1}), and r_a is the aerodynamic resistance to vapor transport (s m^{-1}). These two resistance terms provide a simplified but effective representation of water vapor transport processes near the land surface (Fig. 12-8).
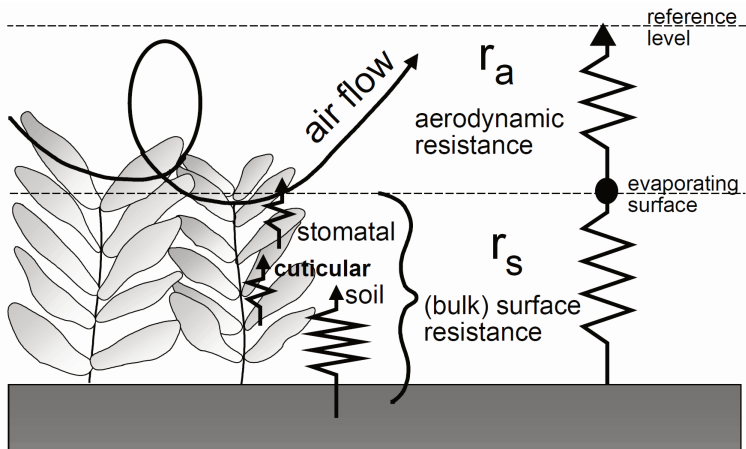
Fig. 12-8. Conceptual representation of near surface water vapor transport as controlled by the surface and aerodynamic resistances. Reproduced from Allen et al. (1998).

A panel of experts from the United Nations Food and Agricultural Organization (FAO) has defined the reference surface as, "A hypothetical reference crop with an assumed height of 0.12 m, a fixed surface resistance of 70 s m^{-1}, and an albedo of 0.23." [8]. Based on this definition, and employing some approximations for the resistance terms, the FAO Penman–Monteith equation for ET$_0$ is:

$$ET_0 = \frac{0.408\Delta(R_n - G) + \gamma\frac{900}{(T+273)}u_2(e_s - e_a)}{\Delta + \gamma(1 + 0.34u_2)}$$

(Eq. 12-7)

where ET_0 is the reference evapotranspiration (mm d^{-1}), R_n and G are in MJ m^{-2} d^{-1}, T is the mean daily air temperature at 2-m height (°C), and u_2 is the mean daily wind speed at 2 m (m s-1). Once ET_0 is determined, site- and day-specific "crop coefficients" (K_c) can be estimated to calculate actual ET from ET_0, i.e. ET = ET_0 × Kc. With judicious use, the FAO Penman-Monteith equation has proven to provide reliable ET_0 estimates for use in crops, grasslands, and forests around the world. For example, the FAO Penman-Monteith equation was used to estimate daily ET_0 during a 3-yr bioenergy crop experiment in Oklahoma (Fig. 12-9). The results show that actual ET can be substantially less than ET_0 during dry periods (e.g. summers of 2011 and 2012) but can exceed ET_0 during wet periods (e.g. summer of 2013) [14].
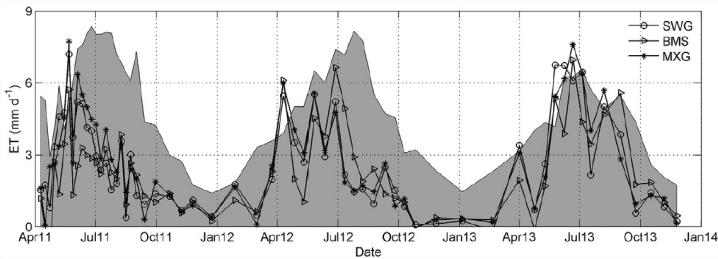
Fig. 12-9. Reference evapotranspiration (shaded region) and actual evapotranspiration under switchgrass (SWG), biomass sorghum (BMS), and mixed perennial grasses (MXG) at Stillwater, Oklahoma. Reproduced from Yimam et al. (2014).

12.6 Problem Set

1. Under daylight conditions, the human eye exhibits maximum sensitivity to light at a wavelength of 555 nm. At what temperature (K) would a blackbody emit light with a peak intensity at this wavelength?

2. The sun has a radius of approximately 6.96 x 108 m and an average surface temperature of 5,780 K. Assume it has an emissivity of 0.990. What is the total radiant flux from the sun in $W\,m^{-2}$ and in W?

3. Imagine the Earth is on the surface of a sphere whose radius is equal to the sun-Earth distance, 1.50 x 1011 m. If the total radiant flux from the sun (W) calculated above is evenly distributed over the surface of that imaginary sphere, what is the average flux in $W\,m^{-2}$?

4. Use the Hargreaves method to estimate monthly ET_0 for Arcadia, Oklahoma, USA in the year 2016, given the data below. Use a spreadsheet and report only the annual total (mm). Assuming a latent heat of vaporization of 2.465 MJ kg^{-1}, one MJ m^{-2} is equivalent to 0.4057 mm of evaporated water.

Table 12-1. Extraterrestrial radiation (Ra), mean temperature ( ), mean of daily maximum temperature ($Tmax$), and mean of daily minimum temperature ($Tmin$) near Arcadia, Oklahoma, USA in 2016. Ra estimates obtained from https://www.nrel.gov/midc/solpos/solpos.html and temperatures obtained from the Oklahoma Mesonet, Spencer site, http://www.mesonet.org/.

Month	R_a	R_a	Tmax	Tmin	T̄
	W m^{-2}	MJ m^{-2}	°C	°C	°C
Jan	210	563	9.52	-1.61	3.66
Feb	270	676	16.1	2.8	9.5
Mar	348	932	19.7	6.6	13.3
Apr	418	1083	22.9	10.5	16.7
May	465	1246	24.8	14.2	19.4
Jun	482	1250	31.7	20.8	26.0
Jul	471	1260	33.6	23.3	28.1
Aug	429	1150	32.9	21.7	26.7
Sep	367	950	30.4	18.1	23.6
Oct	290	778	26.6	13.8	19.9
Nov	223	578	20.0	7.6	13.6
Dec	192	514	9.94	-1.97	4.03

12.7 References

1. Ren, H., et al., *Empirical algorithms to map global broadband emissivities over vegetated surfaces.* IEEE Transactions on Geoscience and Remote Sensing, 2013. **51**(5): p. 2619-2631.

2. Sanchez, J.M., et al., *Thermal infrared emissivity dependence on soil moisture in field conditions.* IEEE Transactions on Geoscience and Remote Sensing, 2011. **49**(11): p. 4652-4659.

3. Ogawa, K., T. Schmugge, and S. Rokugawa, *Estimating Broadband Emissivity of Arid Regions and Its Seasonal Variations Using Thermal Infrared Remote Sensing.* IEEE Transactions on Geoscience and Remote Sensing, 2008. **46**(2): p. 334-343.

4. Cheng, J. and S. Liang, *Estimating the broadband longwave emissivity of global bare soil from the MODIS shortwave albedo product.* Journal of Geophysical Research: Atmospheres, 2014. **119**(2): p. 614-634.

5. Herrero, J. and M. Polo, *Parameterization of atmospheric longwave emissivity in a mountainous site for all sky conditions.* Hydrology and Earth System Sciences, 2012. **16**(9): p. 3139-3147.

6. Staley, D. and G. Jurica, *Effective atmospheric emissivity under clear skies.* Journal of Applied Meteorology, 1972. **11**(2): p. 349-356.

7. Sauer, T.J., et al., *Surface energy balance of a corn residue-covered field.* Agricultural and Forest Meteorology, 1998. **89**(3–4): p. 155-168.

8. Allen, R.G., et al., *Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56, in FAO, Rome.* 1998. p. D05109.

9. Hargreaves, G.H. and R.G. Allen, *History and evaluation of Hargreaves evapotranspiration equation.* Journal of Irrigation and Drainage Engineering, 2003. **129**(1): p. 53-63.

10. Hargreaves, G.H. and Z.A. Samani, *Estimating potential*

evapotranspiration. Journal of the Irrigation and Drainage Division, 1982. **108**(3): p. 225-230.

11. Krueger, E.S., Y.T. Yimam, and T.E. Ochsner, *Human factors were dominant drivers of record low streamflow to a surface water irrigation district in the US southern Great Plains.* Agricultural Water Management, 2017. **185**: p. 93-104.

12. Penman, H.L., *Natural Evaporation from Open Water, Bare Soil and Grass.* Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences, 1948. **193**(1032): p. 120-145.

13. Monteith, J.L. *Evaporation and environment.* Symp. Soc. Exp. Biol. 1965.

14. Yimam, Y.T., et al., *Soil Water Dynamics and Evapotranspiration under Annual and Perennial Bioenergy Crops.* Soil Sci. Soc. Am. J., 2014. **78**(5): p. 1584-1593.
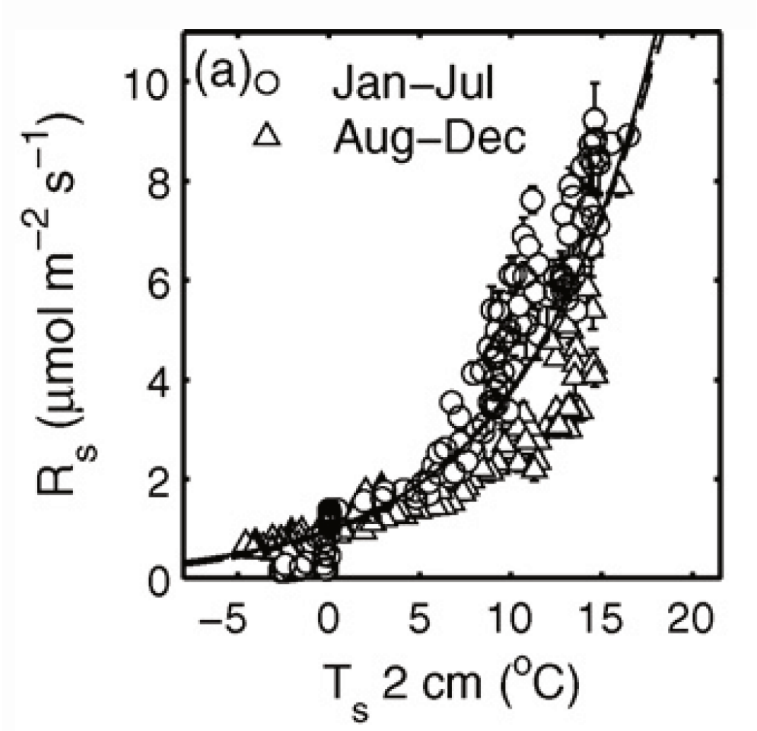
13. SOIL TEMPERATURE

In the prior chapter, we focused on the partitioning of energy at the soil surface between net radiation, latent heat flux, sensible heat flux, and soil heat flux. One of the key variables that controls or is controlled by that energy partitioning is soil temperature. The aim of this chapter is to explain how soil temperatures vary in space and time and how soil temperatures impact a wide variety of biological, chemical, and physical processes in the soil. The global distribution of soil temperature regimes reflects the wide range of soil thermal environments, from soils with a megathermic temperature regime having mean annual temperatures at the 50-cm depth >28°C, to soils in the hypergelic temperature regime with mean annual temperatures <-10°C (Fig. 13-1).
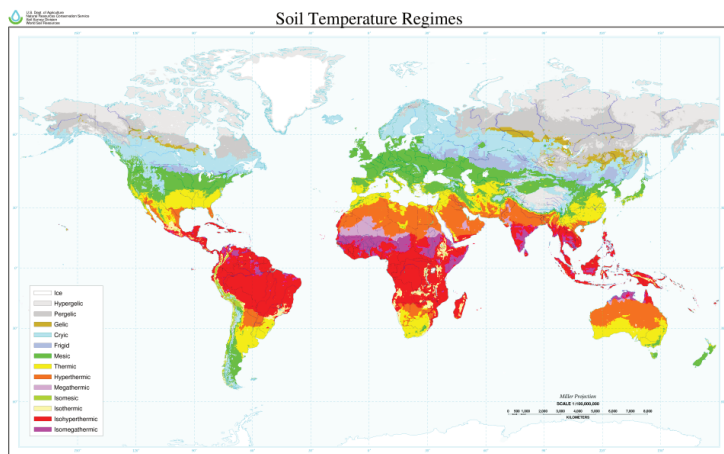
Fig. 13-1. Map of global soil temperature regimes based on an interpolation of over 20,000 climatic stations that were input into a soil water balance model to estimate soil temperature regimes. Created by USDA–NRCS Soil Science Division. Adapted by Ochsner, enlarged legend.

Soil temperature strongly affects the global carbon cycle by influencing the soil respiration rate, which includes respiration of both soil microbes and plant roots. The soil respiration rate under boreal forest in Canada shows a strong logarithmic relationship with soil temperature (Fig. 13-2) [1], and similar relationships have been observed for many other locations and ecosystems around the world. If rising global temperatures cause increased soil respiration resulting in a net transfer of soil organic carbon to the atmosphere, then a positive feedback to climate change would occur [2].

Fig. 13-2. Relationship between daily mean soil respiration (Rs) and daily mean soil temperature at 2 cm (Ts) in boreal forest in Canada. Adapted from Gaumont–Guay et al. (2006).

Multiple plant growth processes can also be strongly influenced by soil temperature. For example, soil temperature is one of the primary factors controlling the germination and early growth of many crops. The time required for winter wheat seedlings to emerge decreases from >25 days at a soil temperature of 5°C to ~5 days at 20°C (Fig. 13-3) [3]. For corn, warmer early season soil temperatures hasten plant development, increase leaf size in the upper half of the canopy, and linearly increase end-of-season grain yields [4]. In recent decades, long-term increases in soil temperature have been documented [5], and these increases may have significance for crop production. Perhaps in response to rising soil temperatures, there has been a trend toward earlier planting dates in the major corn-producing region of the US, the Corn Belt, and these earlier planting dates have contributed significantly to yield increases, especially in the more northern parts of the Corn Belt [6].
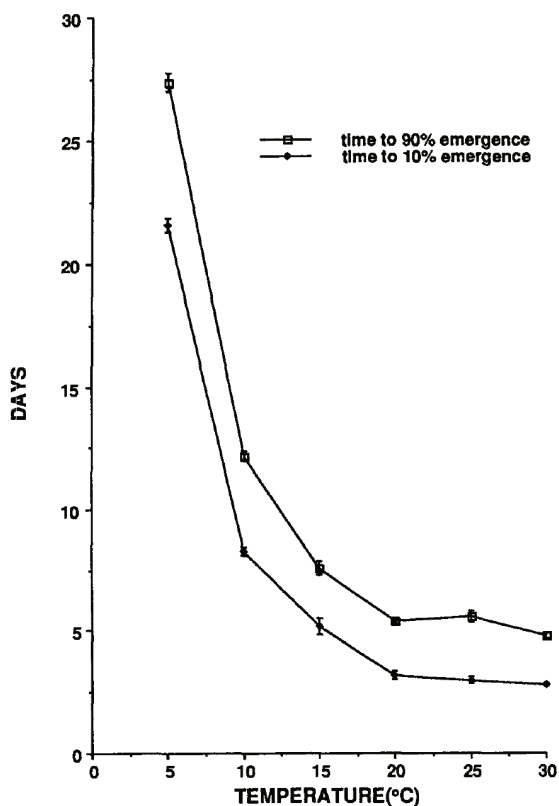
**Fig. 13-3. Influence of soil temperature on time to 10%
emergence and 90% emergence for winter wheat.
Reproduced from Lafond and Fowler (1989).**

13.1 Heat Transfer in Soil

In the soil, heat transfer occurs primarily by conduction, although convective heat transfer can be important in some cases. Heat conduction is governed by Fourier's Law, which was first documented in 1807 and published in 1822 in France and may have influenced the later development of Darcy's Law (1856) [7]. Just as Darcy's Law says that the water flux is proportional to the hydraulic gradient, **Fourier's Law** states that the heat flux by conduction is proportional to the temperature gradient:

$$q_h = -\lambda \frac{dT}{dz}$$

(Eq. 13-1)

where q_h is the heat flux (W m^{-2}), l is the thermal conductivity of the substance through which heat is being conducted (W m^{-1} °C^{-1}), and is the temperature gradient. Fourier's work had tremendous influence, and the basic form of Fourier's Law was followed subsequently in key breakthroughs such as Ohm's Law for electrical conduction and Fick's Law for chemical diffusion, along with Poiseuille's Law and Darcy's Law.

13.2 Soil Thermal Properties

The primary thermal properties of soil, or any substance, are the heat capacity and the thermal conductivity. The heat capacity can be defined per unit mass, in which case it is often called the specific heat, or per unit volume, in which case it is called the volumetric heat capacity. Sometimes it is useful to consider the ratio of the thermal conductivity to the volumetric heat capacity, and this ratio is called the thermal diffusivity. We will define and consider each of these in turn below. Knowledge of the soil thermal properties is necessary to predict how soil temperatures vary in space and time. Sensors which measure soil thermal properties can be used monitor soil water content nondestructively. Soil thermal properties also play a role in several remote-sensing based approaches for estimating soil moisture across large regions.

13.2.1   Thermal Conductivity

The soil ***thermal conductivity*** (λ) is the ratio of the magnitude of the conductive heat flux through the soil to the magnitude of the temperature gradient (W m^{-1} °C^{-1}). It is a measure of the soil's ability to conduct heat, just as the hydraulic conductivity is a measure of the soil's ability to "conduct" water. Soil thermal conductivity is influenced by a wide range of soil characteristics including:
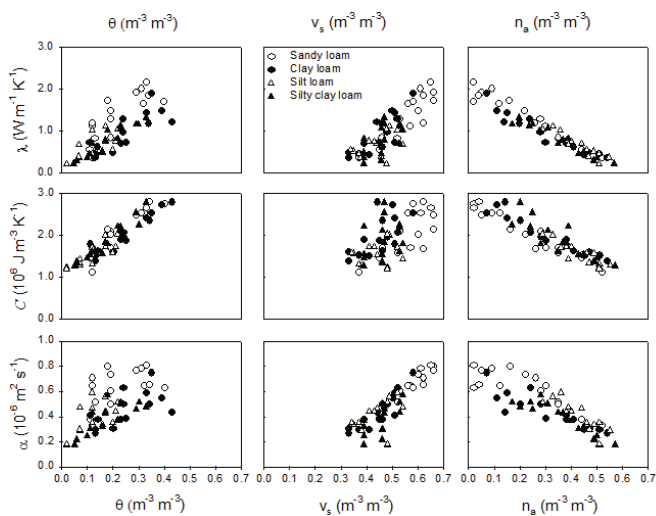
- air-filled porosity
- water content
- bulk density
- texture
- mineralogy
- organic matter content
- soil structure

- soil temperature

Among common soil constituents, quartz has by far the highest thermal conductivity and air has by far the lowest thermal conductivity (Table 13-1) [8] [9]. Often, the majority of the sand-sized fraction in soils is composed primarily of quartz, thus sandy soils have higher thermal conductivity values than other soils, all other things being equal. Since the thermal conductivity of air is so low, air-filled porosity exerts a dominant influence on soil thermal conductivity. The higher the air-filled porosity is, the lower the thermal conductivity is (Fig. 13-4). Soil thermal conductivity increases as water content increases, but not in a purely linear fashion. For dry soil, relatively small increases in the water content can substantially increase the thermal contact between mineral particles because the water adheres to the particles, resulting in a relatively large increase in the thermal conductivity.

Table 13-1. Thermal conductivity, density, and specific heat of common soil constituents at 10 °C (after de Vries, 1963, Table 7.1; adapted from Horton and Ochsner, 2011).

Soil constituent	Thermal conductivity	Density	Specific heat
	W m-1 °C -1	g cm-3	J g-1 °C -1
Quartz	8.8	2.66	0.75
Clay minerals	3	2.65	0.76
Soil organic matter	0.3	1.3	1.9
Water	0.57	1.00	4.18
Ice (0 °C)	2.2	0.92	2.0
Air	0.025	0.00125	1.0

Fig. 13-4. Thermal conductivity (l), volumetric heat capacity (C), and thermal diffusivity (a) as influenced by volumetric water content (q), volume fraction of solids (vs), and air-filled porosity (na) across four different soils. Reproduced from Ochsner et al. (2001).

13.2.2   Heat Capacity

Soil **volumetric heat capacity** (C) is the amount of energy required to raise the temperature of a unit volume of soil by one degree (J m^{-3} °C^{-1}). Unlike thermal conductivity, volumetric heat capacity increases strictly linearly as soil water content increases (Fig. 13-4). Volumetric heat capacity is also a linear function of bulk density. The volumetric heat capacity can be calculated by

$$C = \rho_b c_s + \rho_w c_w \theta$$

(Eq. 13-2)

where ρ_b is the soil bulk density (g cm^{-3}), cs is the specific heat of the soil solids (J g-1 °C-1), ρ_w is the density of water (g cm^{-3}), c_w is the

specific heat of water, and θ is the volumetric water content (cm^3 cm^{-3}). To increase the temperature of wetter, denser soil requires more energy than to increase the temperature of drier, less dense soil, which has a lower volumetric heat capacity. This is one factor that can contribute to lower soil temperatures and delayed crop development in soils managed with no tillage [10].

13.2.3   Thermal Diffusivity

The soil **thermal diffusivity** is the ratio of the thermal conductivity to the volumetric heat capacity (m^2 s^{-1}). It is an indicator of the rate of at which a temperature change will be transmitted through the soil by conduction. When the thermal diffusivity is high, temperature changes are transmitted rapidly through the soil. Logically, soil thermal diffusivity is influenced by all the factors which influence thermal conductivity and heat capacity. Thermal diffusivity is somewhat less sensitive to soil water content than are thermal conductivity and volumetric heat capacity (Fig. 13-4). The thermal diffusivity is a particularly useful parameter to aid in understanding and modeling soil temperatures, which is the next topic we will consider.

13.3 Soil Surface Temperature

The highest and lowest soil temperatures occur at the soil surface. The highest recorded soil temperatures on Earth are near 700°C under an intense forest fire [11]. The lowest recorded soil temperatures on Earth are less extreme, reaching -20 °C in Arctic winters [12]. Regular oscillations in the temperature of the soil surface are driven by the daily cycle of the Earth spinning on its axis and the annual cycle of the Earth orbiting the sun. Irregularities are superimposed on these regular oscillations due to variations in the weather such as the passage of clouds, precipitation events, cold fronts, and warm fronts. If we choose a time period for which such irregularities are absent or negligible, then, as a first approximation, we can represent the soil surface temperature (T_0) using a sine wave:

$$T_0 = T_{avg} + A_0 sin(\omega t + \phi)$$

(Eq. 13-3)

where T_{avg} is the average soil temperature at the surface, A_0 is the amplitude of the temperature oscillation at the surface, w is the angular frequency of the cycle, and f is a phase shift constant. The angular frequency is calculated as , where the period is 24 hours or 365 days. The phase shift is a time offset which is needed to ensure that the maximum and minimum temperatures occur at the appropriate times.
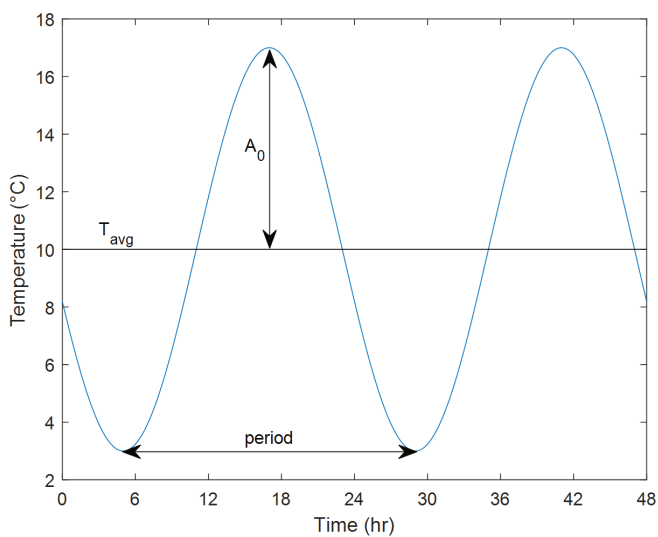
Fig. 13-5. Sinusoidal approximation of soil temperature oscillation with mean temperature of 10 °C, amplitude of 7 °C, period of 24 hours, and phase shift of 3.40 radians or 13 hours.

13.4 Sub-Surface Soil Temperatures

We can also approximate the oscillations of subsurface soil temperatures as a sine wave if we assume:

- the surface temperature is (and has been) oscillating as a sine wave,
- the average soil temperature is the same for all depths, and
- deep in the soil the temperature is constant.

Under these conditions, the soil temperature for any depth (z) and time (t) is given by:

$$T(z,t) = T_{avg} + A_0 e^{-z/d} sin(\omega t + \phi - z/d)$$

(Eq. 13-4)

where e is the base of the natural logarithm and d is called the damping depth.

The **damping depth** is the depth at which the amplitude of the soil temperature oscillation is only 37% ($1/e$) as large as the amplitude at the soil surface. The damping depth depends on the soil thermal diffusivity (α) and the angular frequency (ω) or period of the oscillation.

$$d = \sqrt{2\alpha/\omega} = \sqrt{\alpha \cdot period/\pi}$$

(Eq. 13-5)

The damping depth equation above helps us see that the soil temperature oscillations associated with the annual cycle (period = 365 days) penetrate deeper in the soil than those associated with the daily cycle (period = 1 day). In fact, the annual oscillations penetrate approximately 19.1 times further ( ) than the daily oscillations. The penetration depth also increases as the thermal diffusivity increases.

Plotting Eq. 13-4 helps us to visualize some fundamental aspects of soil temperature dynamics. Figure 13-6 shows the soil temperature dynamics simulated using Eq. 13-4 for transmission of a daily temperature wave in a soil with a thermal diffusivity of 0.50×10^{-6} m^2 s^{-1} (18 cm^2 h^{-1}) and damping depth of 11.7 cm. First, you can see that the amplitude of the soil temperature oscillation decreases exponentially with depth (Fig. 13-6). In this example, the amplitude, which was 7 °C at the soil surface, is only 3 °C by the 10-cm depth. Second, you can see that the maximum (and minimum) temperatures occur progressively later as the soil depth increases. In this example, the soil temperatures at 30 and 40 cm reach their respective maximums during the night or early morning, when the surface temperature is near its minimum. Third, you can see that the depth at which the soil is warmest (or coldest) changes constantly, with each of the plotted depths being the warmest depth in the soil profile for at least a small portion of the day.
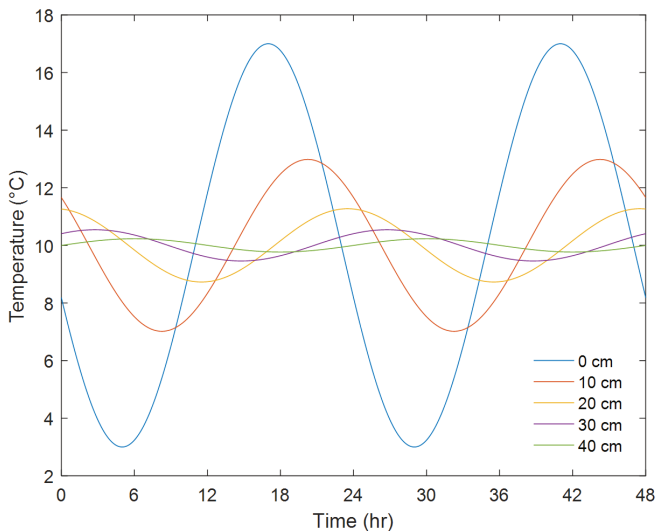

Fig. 13-6. Simulated transmission of a daily soil temperature wave to depths of 0, 10, 20, 30, and 40 cm in a soil with a thermal diffusivity of 0.50×10^{-6} m^2 s^{-1} (18 cm^2 h^{-1}) and damping depth of 11.7 cm.

The time lag that we observed in Fig. 13-6 can be estimated using the following equation:

$$t_2 - t_1 = \frac{z_2 - z_1}{\sqrt{4\pi\alpha/period}}$$

where t_2-t_1 is the time lag for transmission of the temperature wave between depths z_2 and z_1. As the thermal diffusivity increases, the time lag between successive depths decreases. In fact, this equation provides on way of estimating the thermal diffusivity. By recording the times at which daily maximum (or minimum) temperatures occur for two different soil depths, we can estimate the thermal diffusivity for the soil between those depths.
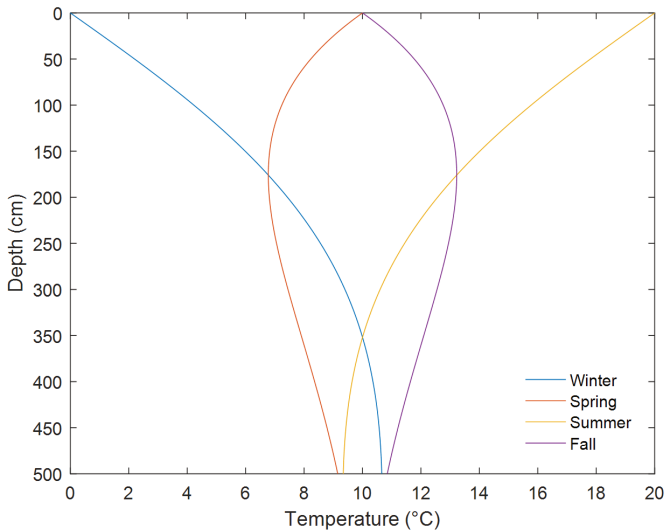
Fig. 13-7. Simulated soil temperature profiles at various times in the year for a soil with average temperature of 10 °C for all depths and an amplitude of 10 °C for the annual oscillation at the soil surface.

We can also use Eq. 13-4 to visualize soil temperature versus depth at defined times. During the winter, the coldest temperature is at the soil surface, and the soil heat flux is upwards towards the surface, i.e. negative (Fig. 13-7). When the soil begins to warm up in the spring, the coldest temperatures are located neither at the surface nor at the bottom of the profile but rather in the middle, in this case near the 175 cm depth. The profiles in the summer and fall are mirror images of the winter and spring profiles, respectively. The same thermal diffusivity value was used for Fig. 13-7 as for Fig. 13-6, but since the former is for annual temperature oscillations the damping depth increases to 224 cm. Keep in mind that these profiles are based on several simplifying assumptions, and they neglect any influences of soil heterogeneity, convective heat transfer, or soil freezing/thawing.
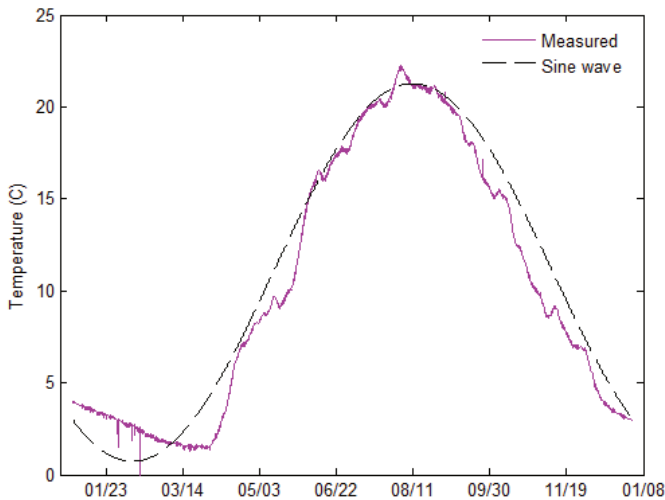
13.5 Measured Soil Temperatures
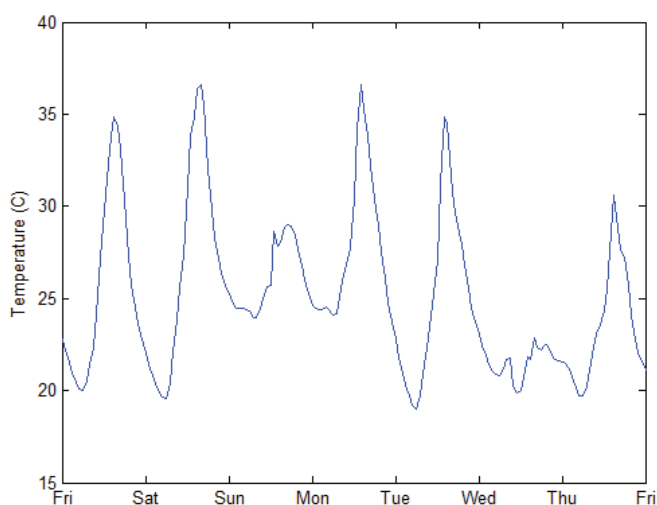
Fig. 13-8. Measured daily oscillations of soil temperature at 6-cm depth in a silt loam soil in southeast Minnesota, USA, under perennial vegetation.

In reality, soil temperature oscillations often bear little resemblance to a regular procession of sine waves. Figure 13-8 shows one week of measured soil temperatures for the 6-cm depth under perennial vegetation in Minnesota, USA. After two days of similar and near-sinusoidal oscillations, the temperatures on the third day reflect a much smaller amplitude and a more irregular temperature pattern. The temperature oscillations on the fourth and fifth days could be reasonably approximated by a sine wave, but on the sixth day again the pattern is disrupted. Soil temperature sequences like this are

not uncommon and typically result from changes in the weather, in this case rainy days in the midst of the growing season.

Fig. 13-9. Measured annual cycle of soil temperature at 1-m depth in a silt loam soil in Minnesota, USA, under perennial vegetation.

As the measurement depth increases, the impacts of short-term weather changes are damped out, and the soil temperature oscillations more frequently reflect a sinusoidal pattern. One year of measured soil temperatures at the 1-m depth for the same site in Minnesota are shown in Fig. 13-9. The data are well-approximated by a sine wave for the majority of the year. The largest deviations from sinusoidal behavior were during the winter and may have been caused by the insulating effects of snow cover on the soil surface and by soil freezing/thawing at depths above 1 m.

13.6 Problem Set

1. The thermal diffusivity for a particular soil is 0.43 x 10-6 m^2 s^{-1}.

   a. What is the damping depth for the diurnal temperature wave?

   b. What is the damping depth for the annual temperature wave?

   c. At what depth is the amplitude of the annual temperature wave only 5% of the amplitude of the annual wave at the surface?

   d. What is the time lag (hr) between the occurrence of the daily maximum temperature at the surface and at 30-cm depth?

13.7 References

1. Gaumont-Guay, D., et al., *Interpreting the dependence of soil respiration on soil temperature and water content in a boreal aspen stand.* Agricultural and Forest Meteorology, 2006. **140**(1–4): p. 220-235.

2. Davidson, E.A. and I.A. Janssens, *Temperature sensitivity of soil carbon decomposition and feedbacks to climate change.* Nature, 2006. **440**(7081): p. 165-173.

3. Lafond, G.P. and D.B. Fowler, *Soil Temperature and Water Content, Seeding Depth, and Simulated Rainfall Effects on Winter Wheat Emergence.* Agronomy Journal, 1989. **81**(4): p. 609-614.

4. Bollero, G.A., D.G. Bullock, and S.E. Hollinger, *Soil Temperature and Planting Date Effects on Corn Yield, Leaf Area, and Plant Development.* Agronomy Journal, 1996. **88**(3): p. 385-390.

5. Baker, J.M. and D.G. Baker, *Long-term ground heat flux and heat storage at a mid-latitude site.* Climatic Change, 2002. **54**: p. 295-303.

6. Kucharik, C.J., *Contribution of Planting Date Trends to Increased Maize Yields in the Central United States.* Agronomy Journal, 2008. **100**(2): p. 328-336.

7. Narasimhan, T., *Fourier's heat conduction equation: History, influence, and connections.* Journal of Earth System Science, 1999. **108**(3): p. 117-148.

8. DeVries, D., *Thermal properties of soils. In 'Physics of plant environment'.(Ed. WR van Wijk) pp. 210–235.* 1963, John Wiley & Sons: New York, NY.

9. Horton, R. and T. Ochsner, *Soil thermal regime*, in *Handbook of Soil Sciences: Properties and Processes, Second Edition.* 2011, CRC Press. p. 1-24.

10. Fortin, M.-C., *Soil Temperature, Soil Water, and No-Till Corn Development Following In-Row Residue Removal.* Agronomy

Journal, 1993. **85**(3): p. 571-576.

11. Neary, D.G., et al., *Fire effects on belowground sustainability: a review and synthesis.* Forest Ecology and Management, 1999. **122**(1-2): p. 51-71.

12. Coulson, S.J., et al., *Thermal Environments of Arctic Soil Organisms during Winter.* Arctic and Alpine Research, 1995. **27**(4): p. 364-370.