A Growing Threat to Farmland: The Geography of Soil Erosion

John P. Wilson

The plentiful land resources of the United States, particularly its highly productive cropland, have played a key role in the nation's economic development and prosperity. As the turn of a new century approaches, however, mounting evidence suggests that accelerating soil erosion may diminish cropland acreage so much that the agricultural sector will be unable to satisfy projected domestic and foreign demands for food and fiber without substantial price increases (Larson 1981). In addition, erosion and non-point source pollution from agriculture may severely limit clean water for drinking, general public use, agricultural and industrial use, recreation, and a generally healthy environment (Schaller and Bailey 1983).

The term "soil erosion" encompasses a group of detachment and transport processes that loosen or dissolve rocky and earthy materials and remove them from the earth's surface through water or wind action. Detaching forces include glacial ice, excessive tillage, wind, flowing water, and crushing by vehicles, animals and people. Transporting forces include glacial ice, gravity, strong wind and running water (Donahue et al. 1983).

Water erosion is the most widespread of soil degradation processes occurring, at least in a minor way, on soils exposed to rainfall, melting snow, furrow irrigation and overland flow. It occurs on sloping land where the intensity of rainfall exceeds the soil's capacity to absorb it, resulting in runoff. The severity of this kind of erosion is determined by five major factors:

- the soil's resistance to breakdown by raindrops or running water, which, in turn, is a function of particle size distribution,
- organic matter content, permeability, level of aggregation and structural stability,
- the intensity of rainfall or runoff events,
- the slope gradient and length, which determine the amount and rate of runoff concentration,
- the presence of frozen or impermeable layers in the soil profile, and
- the vegetation cover or crop residues that protect the soil from raindrop impact and also slow runoff and soil movement (Sparrow et al. 1984).

Wind erosion, on the other hand, begins when particles on the soil surface loosened by the wind bounce along the surface and remove other particles. A less important process of wind erosion is surface creep, which involves the sliding and rolling motion of heavier particles. The most visibly striking action of the wind is the suspension of fine particles in turbulent air, blowing away as clouds of dust. This process was responsible for the great dust storm of May 1934, which originated in western Kansas, Texas, Oklahoma and contiguous portions of Colorado and New Mexico, carrying clouds of powdery debris east to the Atlantic seaboard and hundreds of miles out over the ocean (Brady 1974).
Overall, four factors influence the rate and severity of wind erosion:

- the resistance of soil particles to being moved along the ground by the drag of the wind, a resistance determined by the size of the soil particles and their aggregates and their moisture content,
- the velocity of the wind, which depends partly on the shelter provided by windbreaks and crops,
- the roughness of the soil surface, which determines the drag of the wind on the surface itself, and
- the plants or crop residues on the soil surface, which protect it from the wind (Sparrow et al. 1984).

Wind erosion is primarily a problem in dry regions with fine, sandy soils. Particles larger than fine sand are large enough to resist most winds, and smaller particles tend to clump together and form larger aggregates. However, all soil types are subject to wind erosion if they are dry enough and/or the wind speed is high enough (Sparrow et al. 1984).

These erosion processes were probably often beneficial before the widespread cultivation of crops, leading as they did to the formation of fertile deltas and valleys (Larson et al. 1983). For example, during the latest period of landscape evolution (Wisconsin to recent geologic time), erosion may have been cyclic, occurring as a series of cut-and-fill cycles separated by stable periods (Butzer 1974, Larson et al. 1983).

The problem today is the sustained acceleration of erosion resulting from human activities.

The individual processes of accelerated erosion are similar to those that operate naturally on gentle to intermediate slopes through a broad range of environments. The major difference is that erosion rates are higher, gentler slopes are affected, and a larger variety of erosion processes can be found in one area (Butzer 1974). This accelerated erosion produces both on-site soil and crop damage and off-site damage through air and water pollution. The degree of damage is determined, to a large extent, by the nature of the soil and its position in the landscape.

Soil can be defined as a shallow zone of intermixed mineral and organic matter exhibiting one or more horizons that differ from the underlying regolith in morphology, particle size, chemical composition and biological characteristics. The processes of soil formation include the processes of erosion; a soil survives erosion because it is protected or forms as fast as it dissipates. Figure 1 illustrates the layers or horizons in soil profiles exposed in road cuts and ditch banks and summarizes the characteristic physical and chemical properties for a hypothetical soil profile. Soils with similar horizons are grouped into soil series, and two or more soil series in a particular landscape from a soil association. The letters A, B and C are used to designate the soil horizons in a vertical soil profile (Montagne et al. 1982, Larson et al. 1983).

A horizons, or topsoil, lie at or near the surface and are characterized by maximum accumulation of organic matter and maximum leaching of clay materials, iron and aluminum. The A horizon is the most "active" horizon, where plant roots, soil microorganisms, insects, animals and other forms of life interact with the soil. B horizons, or subsoil, consist of weathered material with maximum accumulations of iron and aluminum oxides and silicate clays. Although B horizons tend to have less biological activity and lower organic matter content than A horizons, they are an important source of nutrients and water for plants, and they normally contain a significant portion of the plant roots. C horizons are made up of unconsolidated material underlying the A and B horizons and are little affected by soil-forming processes. Although biological activity is much lower here, some plant roots penetrate into the C horizons, and it can be an important source of water, particularly after the overlying horizons have dried out (Montagne et al. 1982).

Not all horizons are present in all soils because the interactions of the soil-forming processes produce different soil profiles. For example, because the A horizon forms first, a very young soil may have only an A horizon.

**FIGURE 1**

Hypothetical Soil Profile Showing Major Horizon Symbols and Descriptions (after Larson et al. 1983)

---

John P. Wilson is assistant professor of geography at Montana State University. He holds a BS and an MS in geography and a law degree, all from the University of Canterbury, and a PhD in geography from the University of Toronto. His current interests and activities focus on soil erosion measurements and models and geographic information systems and on the relationship between soil erosion and crop productivity in the northern Great Plains and Ontario.
horizon over a layer of parent material. Alternatively, a heavily eroded soil may have lost all of its A horizon, leaving only the B. Whatever the circumstances, the tendency to develop horizons is a fundamental characteristic of soil formation processes in mineral soils, and the resulting vertical soil profile provides the key to understanding a soil's vulnerability to erosion (Larson et al. 1983).

Some kind of soil profile will maintain itself under most normal circumstances, although its thickness will vary with conditions that reflect climate, vegetation cover, bedrock, slope, length, slope gradient and degree of human interference. On undisturbed moderate and steep slopes, natural erosion is persistent and requires the continuous formation and replacement of soil products. Consequently, most soils on moderate and steep slopes are not inherently stable and remain permanently “young” because they are always composed of fresh products. The disturbance or elimination of vegetation by a variety of human activities may upset this balance so that an established soil is suddenly removed faster than it forms. Agricultural activities are especially relevant in this context because they are so pervasive and because almost all agricultural operations tend to increase soil erosion. Cropland leaves soil bare during critical times — that is, periods of heavy rainfall and snowmelt — while grazing removes at least part of the vegetation throughout the year.

Although concern about accelerated erosion is now widespread, there is no consensus about its extent, effects on crop productivity and the environment, and socioeconomic impacts (Crosson 1983, Larson et al. 1983).

Researchers first started noticing and reporting soil erosion problems in the early 1930s. A national survey to evaluate the severity of the erosion problem was undertaken in 1934, but it was designed to measure the extent to which land was eroded at that point, not how much was occurring annually. The data bases used in subsequent efforts were obscure, inadequate and variable, suggesting little progress in information about soil erosion until the first U.S. Department of Agriculture National Resources Inventory (NRI) was done in 1977 (Larson et al. 1983). This inventory collected data for the Universal Soil Loss Equation (USLE) from a 0.7 percent sample of all non-federal rural lands around the country; for the Wind Erosion Equation (WEE), the survey used a 0.7 percent sample of non-federal rural land in the ten Great Plains states.

These two equations were developed in the 1960s to estimate long-term (10 years or more) water and wind erosion soil losses from cropland under specific types of crop and soil management. The USLE calculates soil losses from water (sheet and rill) erosion as a function of six factors:

\[ A = RKLSCP \]

where A is estimated soil loss (in tons per acre), R is rainfall erosivity (a function of local precipitation characteristics), K is soil erodibility (a function of soil properties), L is slope length, S is slope gradient, C is cover management (crop sequence, crop type, tillage practices and residue management), and P is supporting practices (contouring, strip-cropping or terracing) (Wischmeier and Smith 1978).

The soil losses computed with this equation refer to soil that actually leaves the field or slope, not to in-field sediment movement. It does not predict sediment delivery to streams because the sediment that leaves the field or slope may be deposited at the bottom of the eroded slope and other low areas.

Wind erosion is calculated in the WEE as a function of five factors:

\[ E = f(IKLCV) \]

where E is the potential annual soil loss, f indicates that erosion is a function of the items listed in the equation, I is the soil erodibility index (based on texture and aggregation and analogous to K in the USLE), K is surface roughness, C is climate (wind speed and direction, effective soil moisture), L is the field length (unsheltered distance across a field and parallel to the prevailing wind direction), and V is vegetative cover (Skidmore and Woodruff 1968, Skidmore et al. 1970).

Both equations are used to indicate the probable locations of potential erosion problems and to determine which management practices would achieve the minimum or desired rate of soil loss. They were used in the 1977 inventory to generate the regional and national cropland erosion estimates shown in Figure 2. Results indicate that two-thirds of the erosion is caused by water (sheet and rill erosion) and one-third by wind. Soil loss per hectare exceeded the national average in four of the ten regions classified by the U.S. Department of Agriculture: the Corn Belt (Ohio, Indiana, Illinois, Iowa, Missouri), Appalachia (Virginia, West Virginia, North Carolina, Kentucky, Tennessee), the Mississippi Delta (Arkansas, Mississippi, Louisiana) and the Southern Plains (Texas, Oklahoma).

Almost three quarters of the total erosion occurred in these four regions, with Texas and Iowa leading the individual states. In the two worst USDA regions, the Corn Belt combines a large cropland area (22 percent of the nation's total) with a substantial sheet and rill erosion hazard, while the Southern Plains combine a smaller cropland area (11 percent of the total) with a very substantial wind erosion hazard.

The effects of this erosion on crop pro-
ductivity and water quality depends on the characteristics of the soils and the landscapes in which they occur.

Accelerated soil erosion often reduces crop productivity because it exposes lower soil horizons, which provide less favorable conditions for crop growth. Obviously, soil serves as a medium in which plant roots can take hold, and it also plays a vital role in the provision of the nutrients, water and air essential for plant growth. Some soils do a better job of providing these services than others. The most important soil characteristics influencing plant growth are soil texture, soil structure, organic matter and depth of rooting zone. The precise relationships between these characteristics and crop productivity vary among soils and are not well understood (Crosson 1983).

The A horizons of cultivated soils frequently provide a favorable combination of light, mechanical support, heat, air, water and nutrients. Erosion has reduced or removed these horizons in many cultivated soils. When tilled, the A horizons may be mixed with the upper B horizons, producing a surface soil with texture and other characteristics quite different from those in uneroded conditions. These changes may be important because the stability of the soil surface and the rates at which it conducts water affect the amount of erosion that occurs (Crosson 1983, Larson et al. 1983).

Many soils have B horizons that do not encourage plant root growth. Among these are horizons with excessive accumulations of clay (argillic), high density and strength (fragile), cement-like qualities (duric), low pH (acidic), salt accumulation (salic) and a high aluminum saturation. In addition, the B horizon often controls water permeability. Hence, soils with unfavorable B horizons pose a double threat by increasing the potential for runoff and erosion and by forming a barrier to root development as erosion brings these horizons closer to the surface (Larson et al. 1983).

The USDA has traditionally evaluated these impacts by assigning a soil loss tolerance (T) value to most of the soils mapped in the United States. The T value "denotes the maximum level of soil erosion that will permit a high level of crop productivity to be maintained economically and indefinitely" (Wischmeier and Smith 1978). T values never exceed 11.2 tonnes per hectare annually and are lower with some soils. These values partially explain why soil erosion is so difficult to observe and measure after it has occurred: A soil eroding at the maximum rate will lose less than one centimeter of soil in 10 years (Crosson 1983, Larson et al. 1983).

T values were used in the 1977 NRI to measure the extent and severity of erosion. Overall results showed that water and wind erosion exceeded 11.2 tonnes per hectare on 32 percent of the nation's cropland. Wind erosion by itself exceeded the maximum T value on 51 percent of Texas cropland.

Some researchers argue that damage to the cropland resource because of accelerated erosion is even more severe than the picture painted by this analysis. The concept of the T value is vague, and "establishment of tolerance values has been largely a matter of judgment based on observations" (Wischmeier and Smith 1978).

Several attempts to model the productivity effects of erosion have been made in recent years, but none of them has been used to estimate the national extent and severity of productivity declines. One of the simplest and most promising models involves efforts by Pierce et al. (1983, 1984a, 1984b) and others to estimate the long-term effects of soil erosion on crop production, particularly in terms of irreplaceable soil attributes such as depth, subsoil acidity and water storage capacity. Replaceable attributes, such as plant nutrients and a suitable pH in the tilled layer, can be added at reasonable cost. By using parameters of available water capacity, bulk density, pH and permeabili-
ty and then assigning importance to them in accordance with an idealized root distribution, Pierce et al. (1984a) computed the change in a soil productivity index for the Corn Belt after 25, 50 and 100 years of the erosion rates reported in the 1977 NRI.

The model assumes that soils with favorable characteristics throughout a deep profile exhibit little change in soil productivity due to erosion; soils with unfavorable characteristics in the subsoil or parent material undergo serious reductions in productivity over time. Pierce et al. (1984b) also proposed replacing T values with vulnerability curves that show the productivity index plotted against soil removal. The average slopes of these curves are independent of erosion rates, and they represent the relative vulnerability of a particular soil to long-term erosion losses. Steep curves indicate a soil with productivity that is very vulnerable to erosion, and gentle curves point to soils with productivity that will not diminish very quickly when exposed to erosion.

Application of the model to the Corn Belt indicated that continuation of 1977 erosion rates would reduce crop yields between five and ten percent over 100 years (Pierce et al. 1984b). Crosson (1983) concluded that this and similar estimates suggest that such erosion rates, if continued over the next half century, would present a major threat to the productivity of the nation’s cropland. However, he also pointed out that estimates of crop-yield loss measured in terms of a productivity index or kilograms per hectare lack the economic dimension needed to judge the significance of the loss. The key question is, what effect would such a loss have on the costs of producing crops over the long term?

‡

The extent and character of erosion’s off-site impacts may be more or less serious than on-site impacts, depending on regional landscape characteristics. Larson et al. (1983) noted that most cultivated landscapes fall between two topographic extremes. The first has little relief, no major surface outlet and containment of runoff water and transported sediment in bottomlands. Little or no sediment may leave the cultivated area in these landscapes. The second extreme has distinct slopes and deep, incised stream valleys. Erosion may be severe, and a relatively large amount of the sediment may leave the cultivated area to be deposited in the floodplain or carried farther away (Larson et al. 1983). Both of these situations may cause problems.

Eroded sediment that does not enter a stream and leave the local area may be deposited at the toe of the eroded slope. It may be deposited on cropland, wasteland, pasture or forest land. Occasionally, large amounts of soil may slide downhill to cover homes and roadways.

Eroded sediment that enters streams and lakes may impair water quality, reducing the suitability of water for human consumption and recreational use. Soil itself can become a spectacular pollutant, as is obvious when receding floodwaters leave behind muddy debris, or reservoirs and harbors fill with silt. Donahue et al (1983) estimated that more than 12,300 hectare-meters of sediment settle in reservoirs annually, reducing water storage by the same volume; this amount of water could irrigate 100,000 hectares of alfalfa in the dry areas of the western United States.

More important, the sediments frequently serve as carriers for other materials, notably pathogenic bacteria, nutrients, biocides and heavy metals. These pollutants may upset the ecological balance in aquatic ecosystems by, for example, increasing productivity and thus encouraging accelerated eutrophication of water bodies. Alternatively, the productivity of aquatic ecosystems may be impaired because suspended sediments reduce light penetration, or the sediments interfere with fish spawning and reduce the survival of eggs and fish fry (Pimental et al. 1976). Finally, sediment-laden water is more expensive to treat for...
human use, and deposited sediment is difficult and expensive to remove from agricultural land, residences, roadways, harbors and streams.

These examples illustrate some of the substantial off-site impacts of accelerated soil erosion. In response to this increasingly obvious problem, there have been several efforts during the past 20 years to develop non-point source pollution models to simulate the water quality effects of accelerated soil erosion and sediment delivery to streams and lakes and to provide more precise assessments of these problems. The Field-Scale Model for Chemicals, Runoff and Erosion from Agricultural Management Systems (Knisel 1980, Knisel 1983), the Areal Non-Point Source Watershed Environmental Response Simulation (Beasley et al. 1977, 1982) and the Hydrologic Simulation Program in Fortran (Donigian et al. 1983) are all examples of models that simulate the movement of sediments and other chemicals from fields to streams and lakes at a variety of spatial and temporal scales.

None of these models has been used to define the water quality aspects of cropland erosion on a national scale. Nonetheless, it is apparent that the costs of sedimentation and other kinds of water pollution are substantial; these costs may indeed exceed the expense of holding the soil on the land from which it is eroding.

LITERATURE CITED


Soil accumulation, central Montana

Photo/J.W. Bauder