CHAPTER 1

Digital Terrain Analysis

John P. Wilson and John C. Gallant

1.1 PRINCIPLES AND APPLICATIONS

The development and application of the *TAPES: Terrain Analysis Programs for the Environmental Sciences* software tools described in this book was motivated by our view of the world as a stage on which a series of hierarchically scaled biophysical processes are played out (Figure 1.1). This approach is useful because it can handle the complexity of individual landscape processes and patterns as well as some of the difficulties that are encountered in delineating the appropriate spatial and temporal scales (O'Neill et al. 1986, Mackey 1996, Malanson and Armstrong 1997). Many of the important biophysical processes operating at or near the earth's surface are influenced by both past events and contemporary controls, interactions, and thresholds (Dietrich et al. 1992, Grayson et al. 1993, Montgomery and Dietrich 1995). These interrelationships are complicated and may be best understood using a dynamic systems modeling approach (Kirkby et al. 1996). The boundaries separating different spatial and temporal scales are not very clear and they may vary with individual processes and/or landscapes (cf. Sivapalan and Wood 1986, Mackey 1996, Malanson and Armstrong 1997).

This state of affairs suggests that additional work is required to identify the important spatial and temporal scales and the factors that influence or control the processes and patterns operating at particular scales. The potential benefits may be substantial. Schaffer (1981), working with interacting systems of populations in community ecology, and Phillips (1986), working on examples in fluvial geomorphology, have demonstrated that the key processes operating over different timescales can be considered independently of each other. Phillips (1988) has also shown how the key processes operating at different spatial scales and affecting the hydraulic gradient of a desert stream in Arizona can be considered independently of each other. Band et al. (1991) generated landscape units with low internal variance and high between-unit

Terrain Analysis: Principles and Applications, Edited by John P. Wilson and John C. Gallant.
ISBN 0-471-32188-5

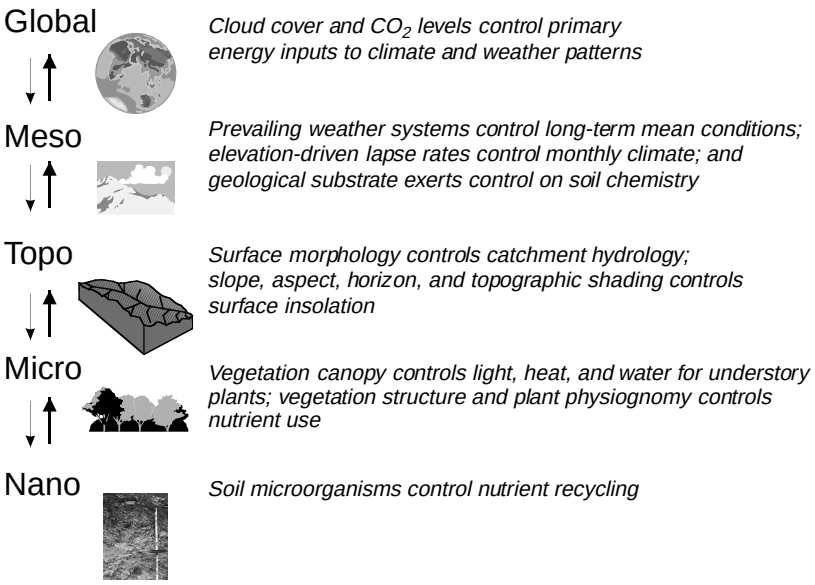

Figure 1.1. Scales at which various biophysical processes dominate calculation of primary environmental regimes. Reprinted with permission from Mackey (1996) The role of GIS and environmental modeling in the conservation of biodiversity. In *Proceedings of the Third International Conference on Integrating GIS and Environmental Modeling, Santa Fe, New Mexico, 21–25 January, 1996,* edited by NCGIA. Copyright © 1996 by National Center for Geographic Information and Analysis, University of California, Santa Barbara.

variance for the important parameters in a nonlinear, deterministic model designed to simulate carbon, water, and nitrogen cycles in a forest ecosystem using a series of hillslope and watershed templates. However, this result may not be universally applicable. Phillips (1988) warned that the key differences in spatial scales cannot be related to fundamental landscape units in numerous instances. Grayson et al. (1993) argued that we should avoid implementing at one scale models developed at a different scale because the simplifying assumptions will often undermine the validity of the original models. Kirkby et al. (1996) concluded that different processes and interactions are likely to emerge as dominant as we move from the plot scale to catchment and regional scales in soil erosion modeling applications. This state of affairs is true of other hydrological, geomorphological, and biological settings as well.

Most of the hydrological, geomorphological, and ecological research of the past century has been conducted at the global and nano- or microscales identified in Figure 1.1 (Mackey 1996). The meso- and toposcales have received much less attention, and yet these scales are important because many of the solutions to environmental problems, such as accelerated soil erosion and non-point-source pollution, will require changes in management strategies at these landscape scales (Moore and Hutchinson 1991). The influence of geologic substrate on soil chemistry (e.g., Likens et al. 1977) and impact of prevailing weather systems and elevation-driven

lapse rates on long-term average monthly climate (e.g., Daly et al. 1994, Hutchinson 1995) exemplify some of the controls operating at the mesoscale. The influence of surface morphology on catchment hydrology and the impact of slope, aspect, and horizon shading on insolation probably represent the most important controls operating at toposcales. Numerous studies have shown how the shape of the land surface can affect the lateral migration and accumulation of water, sediments, and other constituents (e.g., Moore et al. 1988a). These variables, in turn, influence soil development (e.g., Kreznor et al. 1989) and exert a strong influence on the spatial and temporal distributions of the light, heat, water, and mineral nutrients required by photosynthesizing plants (Mackey 1996). The increased popularity of work at these two intermediate scales during the past decade has capitalized on the increasing availability of high-resolution, continuous, digital elevation data and the development of new computerized terrain-analysis tools (Wilson 1996, Burrough and McDonnell 1998, Wilson and Burrough 1999).

1.1.1 Digital Elevation Data Sources and Structures

Most of the currently available digital elevation data sets are the product of photogrammetric data capture (I. D. Moore et al. 1991). These sources rely on the stereoscopic interpretation of aerial photographs or satellite imagery using manual or automatic stereoplotters (Carter 1988, Weibel and Heller 1991). Additional elevation data sets can be acquired by digitizing the contour lines on topographic maps and conducting ground surveys. The advent and widespread use of Global Positioning Systems (GPS) in agriculture and other settings provides many new and affordable opportunities for the collection of large numbers of special-purpose, one-of-a-kind elevation data sets (Fix and Burt 1995, Twigg 1998, Wilson 1999a).

These digital elevation data are usually organized into one of three data structures—(1) regular grids, (2) triangulated irregular networks, and (3) contours—depending on the source and/or preferred method of analysis (Figure 1.2). Square-grid digital elevation models (DEMs) have emerged as the most widely used data structure during the past decade because of their simplicity (i.e., simple elevation matrices that record topological relations between data points implicitly) and ease of computer implementation (I. D. Moore et al. 1991, 1993f, Wise 1998). These advantages offset at least three disadvantages. First, the size of the grid mesh will often affect the storage requirements, computational efficiency, and the quality of the results (Collins and Moon 1981, I. D. Moore et al. 1991). Second, square grids cannot handle abrupt changes in elevation easily and they will often skip important details of the land surface in flat areas (Carter 1988). However, it is worth noting that many of the problems in flat areas occur because the U.S. Geological Survey (USGS) and others persist in recording elevations in whole meters. Third, the computed upslope flow paths will tend to zigzag across the landscape and increase the difficulty of calculating specific catchment areas accurately (Zevenbergen and Thorne 1987, I. D. Moore et al. 1991). Several of these obstacles have been overcome in recent years. For example, there is no generic reason why regular DEMs cannot represent shape well in flat areas, so long as the terrain attributes are calculated by a method

that respects surface drainage. ANUDEM (Hutchinson 1988, 1989b) is one such method and is described in more detail in Chapter 2. Similarly, the advent of several new compression techniques have reduced the storage requirements and improved computational efficiency in recent years (e.g., Kidner and Smith 1992, Smith and Lewis 1994). DEMs with grid sizes of 500, 100, 30, 10, and even 1 m are increasingly available for different parts of the globe (see U.S. Geological Survey 1993, Ordinance Survey 1993, and Hutchinson et al. 1996 for examples).

Triangulated irregular networks (TINs) have also found widespread use (e.g., Tajchman 1981, Jones et al. 1990, Yu et al. 1997). TINs are based on triangular elements (facets) with vertices at the sample points (I. D. Moore et al. 1991). These facets consist of planes joining the three adjacent points in the network and are usually constructed using Delauney triangulation (Weibel and Heller 1991). Lee (1991) compared several methods for building TINs from gridded DEMs. However, the best TINs sample surface-specific points, such as peaks, ridges, and breaks in slope, and form an irregular network of points stored as a set of *x*, *y*, and *z* values together with pointers to their neighbors in the net (I. D. Moore et al. 1991). TINs can easily incorporate discontinuities and may constitute efficient data structures because the density of the triangles can be varied to match the roughness of the terrain (I. D. Moore et al. 1991). This arrangement may cancel out the additional storage that is incurred when the topological relations are computed and recorded explicitly (Kumler 1994).

The third structure incorporates the stream tube concept first proposed by Onstad and Brakensiek (1968) and divides landscapes into small, irregularly shaped polygons (elements) based on contour lines and their orthogonals (Figure 1.2) (O'Loughlin 1986, I. D. Moore et al. 1988a). This structure is used most frequently in hydrological applications because it can reduce complex three-dimensional flow equations into a series of coupled one-dimensional equations in areas of complex terrain (e.g., Moore and Foster 1990, Moore and Grayson 1991, Grayson et al. 1994). Excellent reviews of digital elevation data sources and data structures are presented by Carter (1988), Weibel and Heller (1991), and I. D. Moore et al. (1991).

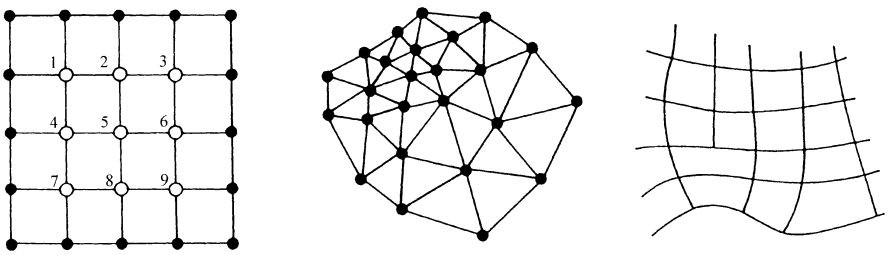

Figure 1.2. Methods of structuring an elevation data network: (a) square-grid network showing a moving 3 by 3 submatrix centered on node 5; (b) triangulated irregular network; and (c) contour-based network. Reprinted with permission from Moore, Grayson, and Ladson (1991) Digital terrain modeling: A review of hydrological, geomorphological, and ecological applications. *Hydrological Processes* 5: 3–30. Copyright © 1991 by John Wiley and Sons Ltd.

The proliferation of digital elevation sources and preprocessing tools means that the initial choice of data structure is not as critical as it once was (Kemp 1997a, b). Numerous methods have been proposed to convert digital elevation data from one structure to another, although care must be exercised with each of these methods to minimize unwanted artifacts (e.g., Krajewski and Gibbs 1994). In addition, larger quantities of data do not necessarily produce better results: Eklundh and Martensson (1995), for example, used ANUDEM (Hutchinson 1988, 1989b) to derive square grids from contours and demonstrated that point sampling produces faster and more accurate square-grid DEMs than the digitizing of contours. Similarly, Wilson et al. (1998) used ANUDEM to derive square grids from irregular point samples and showed that many of the *x, y, z* data points acquired with a truck-mounted GPS were not required to produce satisfactory square-grid DEMs. ANUDEM calculates ridge and streamlines from points of maximum local curvature on contour lines and incorporates a drainage enforcement algorithm that automatically removes spurious sinks or pits in the fitted elevation surface (Hutchinson 1988, 1989b). ANUDEM is one of several programs of this type and an early version has been implemented in the ARC/INFO (Environmental Systems Research Institute, Redlands, CA) geographical information system (GIS) with the TOPOGRID command. Qian et al. (1990) describe an alternative approach that utilizes local operators and global reasoning to automatically extract drainage networks and ridge lines from digital elevation data. Similarly, Smith et al. (1990) proposed a two-step, knowledge-based procedure for extracting channel networks from noisy DEM data. Kumler (1994) described the method used by the U.S. Geological Survey to generate square-grid DEMs from digital contour lines.

Carrara et al. (1997) compared several methods for generating DEMs from contour lines; however, the range of terrain types, sample structures, and modeling routines is so great that attempts to make generalizations about "best" models is tremendously difficult (Burrough and McDonnell 1998, Dixon et al. 1998, Wilson 1999b). In addition, some of the interpolation methods that have been proposed are difficult to use and Eklundh and Martensson (1995) recommended that less experienced users focus on the quality of the input data instead of learning sophisticated interpolation methods. Simpler interpolation methods will give satisfactory results so long as the input data are well sampled and sophisticated algorithms are likely to produce unsatisfactory results if applied to poor data (e.g., Wilson et al. 1998).

1.1.2 Calculation and Use of Topographic Attributes in Hydrological, Geomorphological, and Biological Applications

Many of the most popular topographic attributes, such as slope, specific catchment area, aspect, and plan and profile curvature, can be derived from all three types of elevation data for each and every element as a function of its surroundings (I. D. Moore et al. 1991, 1993f). Individual terrain-analysis tools have been classified in various ways based on the characteristics of the computed attributes and/or their spatial extent. Some authors distinguish tools that perform operations on local neighborhoods (i.e., 3 by 3 moving windows) from those that perform operations on extended neighbor-

hoods (calculation of upslope drainage areas, viewsheds, etc.) (e.g., Burrough and McDonnell 1998). We usually distinguish primary attributes that are computed directly from the DEM and secondary or compound attributes that involve combinations of primary attributes and constitute physically based or empirically derived indices that can characterize the spatial variability of specific processes occurring in the landscape (I. D. Moore et al. 1991, 1993f). This same logic is adopted here.

Primary attributes include slope, aspect, plan and profile curvature, flow-path length, and upslope contributing area (see Table 1.1 for a more complete list). Most of these topographic attributes are calculated from the directional derivatives of a topographic surface. They can be computed directly with a second-order finite difference scheme or by fitting a bivariate interpolation function $z = f(x, y)$ to the DEM and then calculating the derivatives of the function (Moore et al. 1993d, Mitasova et al. 1996, Florinsky 1998). We may or may not want to calculate a depressionless DEM first and we must specify one or more rules to determine drainage directions and the connectivity of individual elements in order to calculate flow-path lengths and upslope contributing areas (e.g., Jenson and Domingue 1988, Martz and De Jong 1988). The overall aim is to be able to use the computed attributes to describe the morphometry, catchment position, and surface attributes of hillslopes and stream channels comprising drainage basins (e.g., Speight 1974, 1980, Band 1986, 1993a, b, Jenson and Domingue 1988, Montgomery and Foufoula-Georgiou 1993, Moore et al. 1993a). Dikau (1989), Dymond et al. (1995), Brabyn (1997), Giles (1998), and Burrough et al. (2000a, b) have all used computed topographic attributes to generate formal landform classifications.

The secondary attributes that are computed from two or more primary attributes are important because they offer an opportunity to describe pattern as a function of process (Table 1.2). Those attributes that quantify the role played by topography in redistributing water in the landscape and in modifying the amount of solar radiation received at the surface have important hydrological, geomorphological, and ecological consequences in many landscapes. These attributes may affect soil characteristics (because the pedogenesis of the soil catena is affected by the way water moves through the environment in many landscapes), distribution and abundance of soil water, susceptibility of landscapes to erosion by water, and the distribution and abundance of flora and fauna. Three sets of compound topographic indices are discussed below to illustrate how these attributes are constructed and used in hydrological, geomorphological, and ecological applications.

Two topographic wetness indices have been used extensively to describe the effects of topography on the location and size of saturated source areas of runoff generation as follows:

$$W_T = \ln\left(\frac{A_s}{T \tan \beta}\right) \tag{1.1}$$

$$W = \ln\left(\frac{A_s}{\tan \beta}\right) \tag{1.2}$$

where A_s is the specific catchment area (m^2m^{-1}), T is the soil transmissivity when the soil profile is saturated, and β is the slope gradient (in degrees) (I. D. Moore et al. 1991,

TABLE 1.1 Primary Topographic Attributes That Can Be Computed by Terrain Analysis from DEM Data

Attribute	Definition	Significance
Altitude	Elevation	Climate, vegetation, potential energy
Upslope height	Mean height of upslope area	Potential energy
Aspect	Slope azimuth	Solar insolation, evapotranspiration, flora and fauna distribution and abundance
Slope	Gradient	Overland and subsurface flow velocity and runoff rate, precipitation, vegetation, geomorphology, soil water content, land capability class
Upslope slope	Mean slope of upslope area	Runoff velocity
Dispersal slope	Mean slope of dispersal area	Rate of soil drainage
Catchment slope	Average slope over the catchment	Time of concentration
Upslope area	Catchment area above a short length of contour	Runoff volume, steady-state runoff rate
Dispersal area	Area downslope from a short length of contour	Soil drainage rate
Catchment area	Area draining to catchment outlet	Runoff volume
Specific catchment area	Upslope area per unit width of contour	Runoff volume, steady-state runoff rate, soil characteristics, soil-water content, geomorphology
Flow path length	Maximum distance of water flow to a point in the catchment	Erosion rates, sediment yield, time of concentration
Upslope length	Mean length of flow paths to a point in the catchment	Flow acceleration, erosion rates
Dispersal length	Distance from a point in the catchment to the outlet	Impedance of soil drainage
Catchment length	Distance from highest point to outlet	Overland flow attenuation
Profile curvature	Slope profile curvature	Flow acceleration, erosion/deposition rate, geomorphology
Plan curvature	Contour curvature	Converging/diverging flow, soil-water content, soil characteristics
Tangential curvature	Plan curvature multiplied by slope	Provides alternative measure of local flow convergence and divergence
Elevation percentile	Proportion of cells in a user-defined circle lower than the center cell	Relative landscape position, flora and fauna distribution and abundance

Source. Adapted with permission from Moore, Grayson, and Ladson (1991) Digital terrain modeling: A review of hydrological, geomorphological, and ecological applications. *Hydrological Processes* 5: 3–30.

TABLE 1.2 Secondary Topographic Attributes That Can Be Computed by Terrain Analysis from DEM Data

Attribute	Definition	Significance
Topographic wetness indices	$W_T = \ln\left(\frac{A_s}{T \tan \beta}\right)$	This equation assumes steady-state conditions and describes the spatial distribution and extent of zones of saturation (i.e., variable source areas) for runoff generation as a function of upslope contributing area, soil transmissivity, and slope gradient.
	$W = \ln\left(\frac{A_s}{\tan \beta}\right)$	This particular equation assumes steady-state conditions and uniform soil properties (i.e., transmissivity is constant throughout the catchment and equal to unity). This pair of equations predicts zones of saturation where A_s is large (typically in converging segments of landscapes), β is small (at base of concave slopes where slope gradient is reduced), and T_i is small (on shallow soils). These conditions are usually encountered along drainage paths and in zones of water concentration in landscapes.
	$W = \ln\left(\frac{A_e}{\tan \beta}\right)$	This quasi-dynamic index substitutes effective drainage area for upslope contributing area and thereby overcomes the limitations of the steady-state assumption used in the first pair of equations.
Stream-power indices	$\text{SPI} = A_s \tan \beta$	Measure of erosive power of flowing water based on assumption that discharge (q) is proportional to specific catchment area (A_s). Predicts net erosion in areas of profile convexity and tangential concavity (flow acceleration and convergence zones) and net deposition in areas of profile concavity (zones of decreasing flow velocity).
	$\text{LS} = (m+1)\left(\frac{A_s}{22.13}\right)^m \left(\frac{\sin \beta}{0.0896}\right)^n$	This sediment transport capacity index was derived from unit stream power theory and is equivalent to the length–slope factor in the Revised Universal Soil Loss Equation in certain circumstances. Another form of this equation is sometimes used to predict locations of net erosion and net deposition areas.
	$\text{CIT} = A_s (\tan \beta)^2$	Variation of stream-power index sometimes used to predict the locations of headwaters of first-order streams (i.e., channel initiation).

Radiation indices	$R_t = (R_{th} - R_{dh})\, F + R_{dh}\, \nu + R_{th}\, (1 - \nu)\, \alpha$	This equation estimates the total short-wave irradiance incident at the earth's surface for some user-defined period ranging in length from 1 day to 1 year. The three main terms account for direct-beam, diffuse, and reflected irradiance. A variety of methods are used by different authors to calculate these individual components. The methods vary tremendously in terms of sophistication, input data, and accuracy.
	$L_{in} = \varepsilon_a \sigma T_a^4\, \nu + (1 - \nu)\, L_{out}$	This equation estimates the incoming or atmospheric long-wave irradiance.
	$L_{out} = \varepsilon_s\, \sigma\, T_s^4$	This equation estimates the outgoing long-wave irradiance.
	$R_n = (1 - \alpha)\, R_t + \varepsilon_s L_{in} - L_{out}$	This equation estimates the net radiation or surface energy budget at the earth's surface for some user-defined period. May or may not account for the effects of clouds depending on the methods and data sources used to estimate individual short-wave radiation components (see Chapter 4 for additional details).
Temperature indices	$T = T_b - \frac{T_{lapse}(Z - Z_b)}{1000} + CS\left(1 - \frac{\mathrm{LAI}}{\mathrm{LAI}_{max}}\right)$	This equation is used to extrapolate minimum air, maximum air, and surface temperatures for a nearby climate station to other parts of the landscape. This equation corrects for elevation via a lapse rate, slope–aspect effects via the short-wave radiation ratio, and vegetation effects via a leaf area index.

1993d). The second equation contains one less term because it assumes uniform soil properties (i.e., that the soil transmissivity is constant throughout the landscape). Wood et al. (1990) have shown that the variation in the topographic component is often far greater than the local variability in soil transmissivity and that Equation 1.2 can be used in place of Equation 1.1 in many landscapes. Both of these indices predict that points lower in the catchment, and particularly those points near the outlets of the main channels, are the wettest points in the catchment, and the soil-water content decreases as the flow lines are retraced upslope to the catchment divide (Wilson and Gallant 1998).

These indices are used in the TOPMODEL (Beven and Kirkby 1979) hydrologic model to characterize the spatial distribution and extent of zones of saturation and variable source areas for runoff generation. O'Loughlin (1986) also used these indices to identify surface saturation zones in landscapes. Burt and Butcher (1986), Jones (1986), and Moore et al. (1988a) used variants of these compound topographic wetness indices to describe the spatial distribution of soil-water content. Moore et al. (1986) showed how the wetness index versus percent saturated source area relationship can be combined with observed stream-flow data and used to estimate the effective transmissivity of a small forested catchment. Sivapalan et al. (1987) used this index to characterize hydrologic similarity, and Phillips (1990) used it to delineate wetlands in a coastal plain drainage basin. Moore et al. (1993b, c) used slope and topographic wetness index to characterize the spatial variability of soil properties for a toposequence in Colorado. Montgomery and Dietrich (1995) used the TOPOG (O'Loughlin 1986) hydrologic model to predict the degree of soil saturation in response to a steady-state rainfall for topographic elements defined by the intersection of contours and stream-tube boundaries. This measure of relative saturation was then used to analyze the stability of each topographic element for the case of cohesionless soils of spatially constant thickness and saturated conductivity in three California, Oregon, and Washington study areas.

These types of static indices must be used carefully to predict the distribution of dynamic phenomena like soil-water content because surface saturation is a threshold process and because of hysteretic effects (Burt and Butcher 1986, I. D. Moore et al. 1991). In addition, there are several important and implicit assumptions in the derivation of the two wetness indices described above. Most notably, the gradient of the piezometric head, which dictates the direction of subsurface flow, is assumed to be parallel to the land surface and steady-state conditions are assumed to apply (Moore et al. 1993d). Several authors have described the pitfalls of using these indices in inappropriate ways. Jones (1986, 1987), for example, discussed the advantages and limitations of wetness indices as indicators of spatial patterns of soil-water content and drainage. Quinn et al. (1995) summarized various problems and described how steady-state topographic wetness indices can be calculated and used effectively in the TOPMODEL hydrologic modeling framework.

In an attempt to overcome the limitations of the steady-state assumption, Barling (1992) proposed a quasi-dynamic topographic wetness index of the form

$$W = \ln\left(\frac{A_e}{\tan\beta}\right) \tag{1.3}$$

where A_e is the effective specific catchment area. Barling et al. (1994) calculated steady-state and quasi-dynamic indices for a catchment near Wagga Wagga in Australia and found that only the quasi-dynamic index correctly predicted that the topographic hollows and not the drainage channels themselves determined the hydrologic response of the catchment. Wood et al. (1997) proposed an alternative index of saturated zone thickness incorporating both spatial and temporal variation in recharge. However, the suitability of these methods as generally applicable tools has yet to be demonstrated.

Several terrain-based stream-power and sediment transport capacity indices have also been proposed (Table 1.2). Stream power is the time rate of energy expenditure and has been used extensively in studies of erosion, sediment transport, and geomorphology as a measure of the erosive power of flowing water (I. D. Moore et al. 1991). It is usually computed as

$$\Omega = \rho g q \tan \beta \tag{1.4}$$

where ρg is the unit weight of water, q is the discharge per unit width, and β is the slope gradient (in degrees). The compound topographic index $A_s \tan \beta$ is, therefore, a measure of stream power, since ρg is essentially constant and q is often assumed to be proportional to A_s. Several researchers have used variations of this index to predict the locations of ephemeral gullies. Thorne et al. (1986) multiplied this index by plan curvature and predicted both the locations of ephemeral gullies and the cross-sectional areas of the gullies after 1 year of development with variants of this new index. Moore et al. (1988a) showed that ephemeral gullies formed where $W > 6.8$ and $A_s \tan \beta > 18$ for a small semiarid catchment in Australia; and Srivastava and Moore (1989) found that ephemeral gullies formed where $W > 8.3$ and $A_s \tan \beta > 18$ on a small catchment in Antigua. I. D. Moore et al. (1991) concluded that threshold values of these indices are likely to vary from place to place because of differences in soil properties. Moore and Nieber (1989) used the stream-power index to identify places where soil conservation measures that reduce the erosive effects of concentrated flow, such as grassed waterways, should be installed. Montgomery and Dietrich (1989, 1992) and Montgomery and Foufoula-Georgiou (1993) used a variation of this index, $A_s (\tan \beta)^2$, to predict the headwaters of first-order streams (i.e., the locations of channel initiation).

A second compound index was derived by Moore and Burch (1986a–c) from unit stream-power theory and a variant used in place of the length–slope factor in the Revised Universal Soil Loss Equation (RUSLE) for slope lengths < 100 m and slopes < 14° as follows:

$$\mathrm{LS} = (m + 1)\left(\frac{A_s}{22.13}\right)^m (\sin \beta / 0.0896)^n \tag{1.5}$$

where $m = 0.4$ and $n = 1.3$ (Moore and Wilson 1992, 1994). Both this and the next equation are nonlinear functions of slope and specific discharge. This new index calculates a spatially distributed sediment transport capacity and may be better suited to

landscape assessments of erosion than the original empirical equation because it explicitly accounts for flow convergence and divergence (Moore and Wilson 1992, Desmet and Govers 1996b).

Another terrain-based sediment transport capacity index can be used to differentiate net erosion and net deposition areas:

$$\Delta T_{cj} = \phi\, A_{sj-}^{m} (\sin \beta_{j-})^{n} - A_{sj-}^{m} (\sin \beta_{j})^{n} \tag{1.6}$$

where ϕ is a constant, subscript j signifies the outlet of cell j, and subscript $j-$ signifies the inlet to cell j (Moore and Wilson 1992, 1994). This index will predict erosion in areas experiencing an increase in sediment transport capacity and deposition in areas experiencing a decrease in sediment transport capacity. Mitasova et al. (1996) implemented variants of these equations in the GRASS (U.S. Army Corps Engineers 1987) GIS. Net erosion areas coincided with areas of profile convexity and tangential concavity (flow acceleration and convergence), and net deposition areas coincided with areas of profile concavity (decreasing flow velocity). These patterns match those observed by Martz and De Jong (1987), Foster (1990), Sutherland (1991), and Busacca et al. (1993) in a variety of landscapes.

Several authors have criticized the use of Equation 1.5 in place of the original slope gradient and length terms in the Revised Universal Soil Loss Equation (Renard et al. 1991) and its predecessors. Interested readers should consult Moore and Wilson (1992, 1994), Foster (1994), Mitasova et al. (1996, 1997), and Desmet and Govers (1996b, 1997) for additional details. Wilson and Lorang (1999) recently summarized the key elements of this debate, and why the terrain-based approach of Mitas et al. (1996) probably represents a superior approach for simulating the impact of complex terrain and various soil and land cover changes on the spatial distribution of soil erosion and deposition.

The third and final set of compound indices is used to estimate the spatial and temporal distribution of solar radiation at the earth's surface. Topography may exert a large impact on the amount of solar energy incident at a location on the earth's surface (Moore et al. 1993f, Dubayah and Rich 1995). Variations in elevation, slope, aspect, and local topographic horizon can cause substantial differences in solar radiation and thereby affect such biophysical processes as air and soil heating, evapotranspiration, and primary production (Gates 1980, Linacre 1992, Dubayah 1992, 1994, Dubayah and Rich 1995). These processes may, in turn, affect the distribution and abundance of flora and fauna. Moore et al. (1993e), for example, used computed radiation and temperature indices to characterize the fine-scale environmental heterogeneity and environmental domains of the five major subalpine forest types for a 20-km^2 study area in the Brindabella Range in southeastern Australia. Hutchins et al. (1976), Kirkpatrick and Numez (1980), Tajchman and Lacey (1986), Austin et al. (1983, 1984), and Noguchi (1992a, b) have also shown that the distributions of solar radiation and vegetation are highly correlated.

Numerous approaches have been proposed to calculate the radiation fluxes and temperature indices used in these types of applications. Most of the radiation models

incorporate one or more of the following equations. The net radiation, R_n, received by an inclined surface can be written as

$$R_n = (1 - \alpha)\,(R_{dir} + R_{dif} + R_{ref}) + \varepsilon_s\, L_{in} - L_{out} = (1 - \alpha)\, R_t + L_n \tag{1.7}$$

where α is the surface albedo, ε_s is the surface emissivity, R_{dir}, R_{dif}, and R_{ref} are the direct, diffuse, and reflected short-wave irradiance, respectively, for which $R_t = R_{dir} + R_{dif} + R_{ref}$, the global short-wave irradiance, L_{in} is the incoming or atmospheric long-wave irradiance, L_{out} is the outgoing or surface long-wave irradiance, for which $\varepsilon_s L_{in} - L_{out} = L_n$, the net long-wave irradiance.

The total short-wave irradiance is estimated by

$$R_t = (R_{th} - R_{dh})\, F + R_{dh}\, \nu + R_{th}\, (1 - \nu)\alpha \tag{1.8}$$

where R_{th} and R_{dh} are the total and diffuse radiation on a horizontal surface, and F is the potential solar radiation ratio ($= R_o/R_{oh}$), which is the ratio of the potential solar radiation (R_o) on a sloping surface to that on a horizontal surface (R_{oh}), and ν is the skyview factor, which is the fraction of the sky that can be seen by the sloping surface. The total and diffuse short-wave irradiances on a horizontal surface are often expressed as functions of the total and diffuse transmittances of the atmosphere and the potential solar radiation on a horizontal surface. These transmittances are functions of the thickness and composition of the atmosphere, such as the water vapor, dust, and aerosol content (Lee 1978, Gates 1980).

The long-wave irradiance components are approximated on a cell-by-cell basis using

$$L_{out} = \varepsilon_s\, \sigma T_s^4 \tag{1.9}$$

$$L_{in} = \varepsilon_a \sigma T_a^4\, \nu + (1 - \nu)\, L_{out} \tag{1.10}$$

where ε_a is the atmospheric emissivity (a function of air temperature, vapor pressure, and cloudiness), σ is the Stefan–Boltzman constant, T_s is the mean surface temperature, and T_a is the mean air temperature. An equation that utilizes modifications of a simple approach proposed by Running et al. (1987), Hungerford et al. (1989), and Running (1991) for estimating the spatial distribution of minimum, maximum, and average air temperature is summarized in Table 1.2.

The different indices that have been proposed vary in terms of the methods, data sources, and assumptions used to estimate individual components. Moore et al. (1993e) developed an approximate method for estimating each of the above fluxes at any location in a topographically heterogeneous landscape in one of these applications. The variation in the potential solar radiation was estimated over a catchment as a function of slope, aspect, topographic shading, and time of year, and then adjusted for cloud, atmospheric, and land cover effects. The variables that serve as model inputs, such as albedo, cloudiness, emissivity, sunshine fraction, mean air and sur-

face temperatures, and clear-sky transmittances, can be varied on a monthly or annual basis (Wilson and Gallant 1999). Hetrick et al. (1993a, b) developed the SOLARFLUX model in the GRID module of the ARC/INFO GIS and used it with latitude, atmospheric transmissivity, slope, aspect, topographic shading, and time of year to estimate direct and diffuse irradiance at each grid point. The effects of cloud cover, which are likely to be substantial in many humid environments, were not accounted for in this model. Kumar et al. (1997) chose a simpler approach and used latitude and a series of topographic attributes derived from a square-grid DEM to estimate clear-sky, direct-beam, short-wave radiation. These relationships are generally straightforward and numerous authors have summarized the appropriate equations for both horizontal and sloping sites (see Lee 1978, Gates 1980, Iqbal 1983, Linacre 1992 for additional details). Most of the challenges (problems) are encountered when atmospheric effects (precipitable water, dust, etc.), cloud cover, and land surface characteristics (albedo) are considered. Dubayah and Rich (1995) have reviewed many of the important computational challenges and errors that are likely to be encountered in building accurate, physically based topographic solar radiation models. Many of their insights are derived from their work combining the Atmospheric and Topographic (ATM) model (Dubayah 1992), ground measurements, and satellite imagery in the Konza Prairie and Rio Grande Basin (Dubayah 1992, 1994; Dubayah and van Katwijk 1992).

Although it is not always apparent to users of terrain analysis, the three sets of indices described above are simplified process models and are not applicable in all situations (Wilson and Gallant 1998). The topographic wetness index, for example, is based on the assumption that the soil hydraulic conductivity decreases exponentially with depth so that subsurface flow is confined to a shallow layer. If this is not the case, the steady-state and quasi-dynamic topographic wetness indices will be poor predictors of the spatial distribution of soil water. An alternative index might be developed to better represent the topographic effect on water distribution, perhaps based on groundwater potential expressed as a simple elevation difference above a local mean or minimum (e.g., Hinton et al. 1993).

The topographic indices introduced on the preceding pages account for the component of the spatial variability of processes that is due to topographic effects. Other spatially variable factors are usually involved, such as soil hydraulic properties and vegetation in the case of soil water. In some instances, the spatial variations in these other attributes are themselves linked to the topographic indices. The spatial variability of soil properties is one case where significant links have been established (e.g., Moore et al. 1993f, Wilson et al. 1994). There are other properties though where explicit incorporation of the spatial variation of other important components of process models would substantially improve the predictive accuracy of topographic indices particularly when working at a broad landscape scale as opposed to the small catchment scale. Surficial geology and, in some cases, climate are likely candidates for inclusion in these types of applications. Some of the applications reported in this book make use of such additional information.

Additional problems may be encountered by the terrain analyst or user because the spatial and statistical distributions of the computed primary and secondary topographic

attributes may be affected by the presence of errors in the source data and the choice of computer algorithm and/or element size (i.e., grid spacing). These aspects are often interconnected, although the review that follows treats each problem separately.

1.1.3 Identification and Treatment of Error and Uncertainty

Nonsystematic and systematic errors in DEMs may confound the expected relationships between computed terrain attributes and terrain-controlled site conditions. These problems may be amplified when first- and second-order derivatives, such as slope and convexity, are calculated (e.g., Bolstad and Stowe 1994). The most serious problems are usually encountered when secondary attributes are derived: The topographic wetness and sediment transport capacity indices are very sensitive to the presence of errors in source (elevation) data in flat areas and to the choice of flow routing algorithm (Moore et al. 1993f).

Many studies have examined the causes, detection, visualization, and correction of DEM errors. Carter (1988), Weibel and Heller (1991) and Kumler (1994), for example, describe the causes of errors in DEMs compiled by different methods. Several methods have been proposed for the detection of errors and estimation of the magnitude and/or spatial distribution of errors (e.g., Polidori et al. 1991, Brown and Bara 1994, Felicisimo 1994, Fryer et al. 1994, Li 1994, Garbrecht and Starks 1995, Lopez 1997). Most of the quantitative estimates have used topographic map elements (e.g., Evans 1980, Skidmore 1989), field measurements (e.g., Bolstad and Stowe 1994, Hammer et al. 1995, Giles and Franklin 1996), or hypothetical (imaginary) DEMs (e.g., Chang and Tsai 1991, Carter 1992, Hodgson 1995) as reference values for these assessments. Most assessments have also examined specific DEM products (e.g., Sasowsky et al. 1992, Brown and Bara 1994), although a few have compared two or more products (e.g., Bolstad and Stowe 1994, Hammer et al. 1995). Other researchers have focused on the development of methods for the visualization (e.g., Kraus 1994, Hunter and Goodchild 1995, 1996, McCullagh 1988) and correction of errors (e.g., Hannah 1981, O'Callaghan and Mark 1984, Jenson and Domingue 1988, Brown and Bara 1994). Several recent studies are discussed in more detail below to illustrate the key issues that have been addressed.

The horizontal and vertical resolution of most square-grid DEMs is such that flow lines become trapped in pits and depressions in key parts of the landscape. Guercio and Soccodato (1996), Jenson and Domingue (1988), Hutchinson (1988, 1989b), Martz and Garbrecht (1998), and Reiger (1998) have all proposed methods for correcting DEMs and/or avoiding these problems. Topographic attributes are computed for depressionless DEMs in many (most) applications that rely on published DEM data sets (as noted earlier).

Brown and Bara (1994) used semivariograms and fractals to detect the presence of errors in 7.5′ USGS 30-m DEMs and evaluated several types of filters for reducing the magnitude of these errors. Their method does not require reference values. It identified the anisotropic conditions (i.e., where the variation in one direction is different from the variation in another direction) that are consistent with the "banding" or "striping" remnants produced by the aerial photograph scanning procedures used

in the production of many USGS DEMs (U.S. Geological Survey 1993). Brown and Bara (1994) also showed how the semivariance and fractal dimensions provided a quantitative basis for applying corrections to mitigate the severity of these problems for a high-relief study area in Glacier National Park, Montana. The anisotropic conditions were greater for slope and curvature than for elevation.

The use of these types of methods for correcting DEMs may cause additional problems. Garbrecht and Starks (1995), for example, reported similar errors for a low relief drainage basin in Nebraska and concluded that the 7.5′ USGS 30-m DEMs could not be used for wetland drainage analysis in these types of landscapes. However, these authors rejected subsequent manipulation of the original USGS DEMs to remove striping because of the additional data degradation that would have occurred. There is no easy way to distinguish real terrain features and errors when applying correction methods like those of Brown and Bara (1994) in many types of landscapes. Garbrecht and Starks (1995) therefore generated a completely new square-grid DEM based on interpolation from digital contours in order to conduct the drainage and wetland analysis tasks at hand.

Hammer et al. (1995) compared field measured and computer-generated slopes. They divided two 16-ha sites in Atchison County, Missouri into 10-m grid cells and field measured slope for each grid cell. The field-measured slope class maps served as templates for cell-by-cell comparisons with computer-generated slope class maps derived from 10- and 30-m DEMs and a standard 1:24,000-scale USDA soil survey. More than 50% of the areas were classified into correct slope classes with 10-m DEM maps. Two iterations of low-pass filters increased the accuracy of these particular maps. The 30-m DEM maps were 30 and 21% correctly classified in the two areas, and the soil survey maps correctly classified >30% of each area but did not capture the fine-scale landscape heterogeneity. The DEM-derived maps underestimated slopes on convexities and overestimated slopes on concavities at both sites.

There are at least three sets of problems connected with the above approach. First, the use of reference values for assessments of multiple attributes over large areas is impractical (Ruiz 1997). Second, Li (1991) and Kumler (1994) both examined the impact of the number, distribution, and accuracy of the checkpoints used for experimental tests of DEM accuracy on the resulting accuracy estimates, and found that the results were often highly sensitive to the choice of reference values. Third, some researchers have argued that the accuracy of primary and secondary topographic attributes cannot be determined by a comparison of calculated and "reference" values because the land surface is not mathematically smooth and there are no actual values of any attributes except elevation (e.g., Shary 1991).

Florinsky (1998) has argued that the accuracy of these attributes depends on the accuracy of the initial data (i.e., the DEM) and the precision of the calculation technique(s). His novel approach incorporated three steps and avoided all three sets of problems noted above. First, he showed that the Evans (1980) method was the most precise method for computing four local topographic variables (i.e., slope, aspect, and plan and profile curvature). The other methods tried were those of Zevenbergen and Thorne (1987), Moore et al. (1993b), and Shary (1995). Second, he derived formulae to calculate root mean square errors (RMSEs) based on the partial derivatives

of the elevation surface for these variables (provided they are estimated with the Evans method). Third, Florinsky (1998) argued that mapping is the most convenient and practical way to implement the formulae that were derived (Figure 1.3). Overall, four general observations about errors can be deduced from this approach:

1. The values of the slope, aspect, and plan and profile curvature RMSEs are directly proportional to the elevation RMSE.
2. The values of these RMSEs increase with decreasing grid spacing.
3. The values of the plan and profile curvature RMSEs are more responsive than slope gradient and aspect to changes in grid spacing.
4. The values of all four RMSEs can become large with decreasing slope gradients (i.e., in flat areas).

These observations also highlight the important roles played by computation methods and grid spacing (grid resolution) in the identification and treatment of error.

Many researchers, including Florinsky (1998), have examined the sensitivity of computed attributes to the method of computation. Gao (1994), for example, proposed a C program for computing slope, aspect, and plan and profile curvature that included special rules for handling edge cells and special areas (summits, ridge lines, stream lines, etc.). Most terrain-analysis methods incorporate special rules for handling these cases. Srinivasen and Engel (1991) compared the performance of four slope algorithms with topographic map and field assessments of slope steepness. Moore et al. (1993d) showed that the D8 slope algorithm (described in more detail below) predicted slightly larger slopes than the finite difference method for a forested study area in southeastern Australia. Weih and Smith (1997) examined the influence of several cell slope computation algorithms on a common forest management decision.

At least six algorithms have also been proposed for routing flow and computing contributing areas from square-grid DEMs. Five of these algorithms—the D8 (deterministic eight-node) algorithm of O'Callaghan and Mark (1984), the Rho8 (random eight-node) algorithm of Fairfield and Leymarie (1991), the FD8 and FRho8 algorithms, and the DEMON algorithm of Costa-Cabral and Burges (1994)—have been implemented in TAPES-G (Moore 1992, Wilson and Gallant 1998). The sixth method uses a vector–grid approach and has been implemented as the *r.flow* routine in the GRASS GIS (Mitasova and Hofierka 1993, Mitasova et al. 1995, 1996).

The D8 algorithm allows flow to one of only eight neighbors based on the direction of steepest descent. This popular algorithm is often criticized because it tends to predict flow in parallel lines along preferred directions that will agree with aspect only when aspect is a multiple of 45° and it cannot model flow dispersion (e.g., Moore et al. 1993d). Rho8 is a stochastic version of D8 that simulates more realistic flow networks but still cannot model flow dispersion. Moore et al. (1993d) found that Rho8 breaks up the long, linear flow paths produced by the D8 method while generating more single-cell drainage areas. Both FD8 and FRho8 allow flow to be distributed to multiple nearest-neighbor nodes in upland areas above defined channels and use either the D8 or Rho8 algorithms below points of channel initiation (Moore et al.

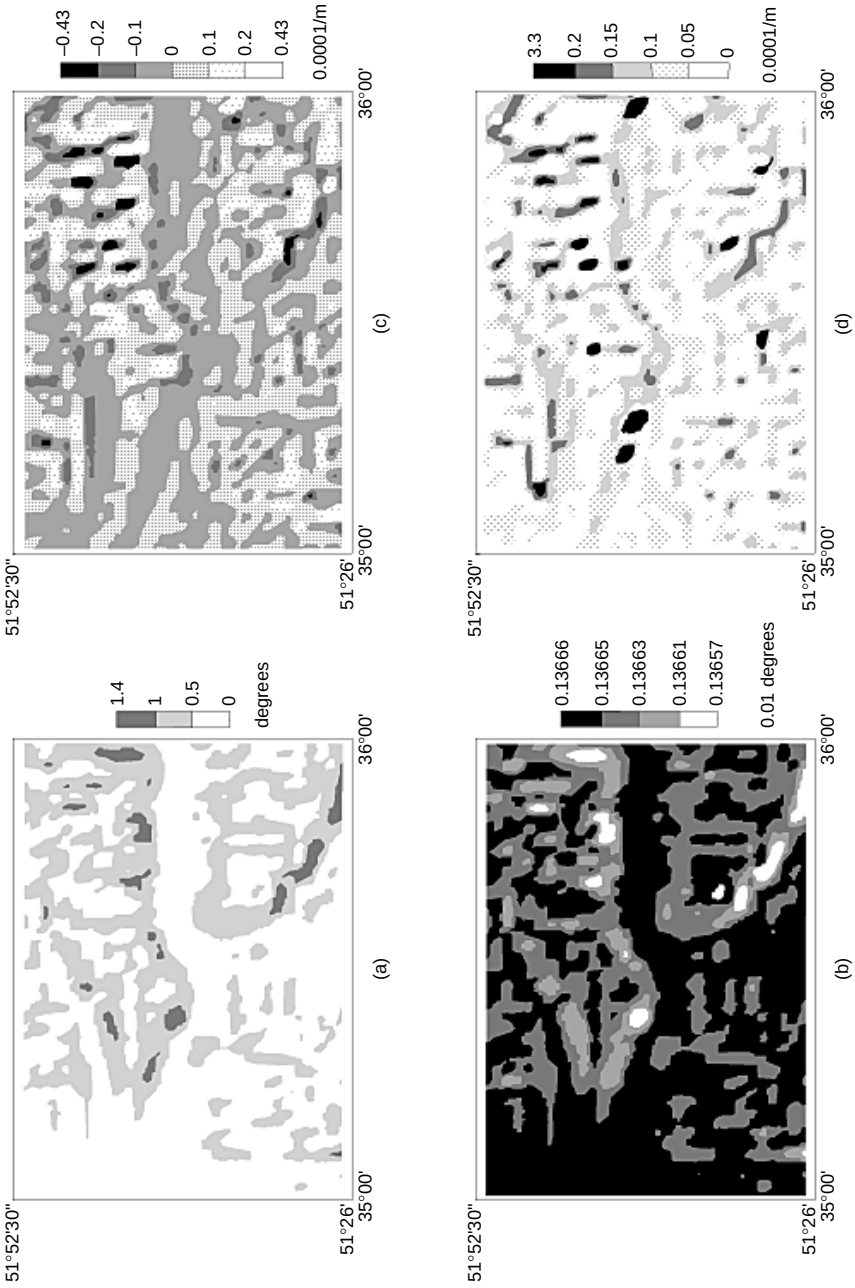

Figure 1.3. Maps of Kursk Region in Russia showing (a) slope gradient; (b) RMSE of gradient; (c) profile curvature; and (d) RMSE of profile curvature. Reprinted with permission from Florinsky (1998) Accuracy of local topographic variables derived from digital elevation models. *International Journal of Geographical Information Science* 12: 47–61 (http://www.tandf.co.uk/journals). Copyright © 1998 by Taylor and Francis.

1993d). Flow in upland areas is assigned to multiple downstream nearest neighbors with these algorithms in TAPES-G using slope-weighted methods similar to those of Freeman (1991) and Quinn et al. (1991). Moore et al. (1993d) showed that the FD8 and FRho8 algorithms implemented in TAPES-G produced almost identical catchment area frequency distributions with very few single-cell drainages. DEMON avoids these problems by representing flow in two directions as directed by aspect. This approach permits the representation of varying flow width over nonplanar topography (similar to contour-based models) (Moore 1996). The final vector–grid method constructs flow lines downhill from each grid cell until they reach a cell with a slope lower than some specified minimum, a boundary line, or some other barrier to calculate upslope contributing areas. These flow lines follow the aspect direction of flow and they are represented in vector format, avoiding the artificial nature of cell-to-cell flow routing in the previous methods. The points defining the flow lines are computed as the points of intersection of a line constructed in the flow direction given by the aspect angle and a grid cell edge. The DEMON and vector–grid algorithms share many similarities and are likely to produce similar results (Wilson and Lorang 1999).

Wolock and McCabe (1995), Moore (1996), Desmet and Govers (1996a), and the two case studies discussed in Chapter 5 provide detailed assessments of the performance of many of the existing algorithms. Wolock and McCabe (1995) compared several single- and multiple-flow-direction algorithms for calculating the topographic parameters used in TOPMODEL. Moore (1996) compared the D8, Rho8, FD8/FRho8, DEMON, and contour-based algorithms in terms of specific catchment area calculations. Desmet and Govers (1996a) compared six flow-routing algorithms in terms of contributing area calculations and the prediction of ephemeral gully locations. These comparisons showed that the single-flow (D8, Rho8) and multiple-flow (FD8, FRho8) direction algorithms will perform very differently in most types of landscapes. The identification of problems with existing algorithms and fundamental role of flowing water in controlling or explaining many key environmental processes and patterns are likely to promote further methodological innovation in this area. Holmgren (1994) and Quinn et al. (1995), for example, have recently proposed new methods for computing the weights used with the FD8 and/or FRho8 algorithms. Burrough et al. (1999a, b) introduced random errors and computed several hundred realizations of the flow network with the D8 algorithm to overcome the presence of DEM errors and/or shortcomings of this algorithm noted earlier.

An even larger group of studies have examined the sensitivity of selected attributes to the choice of data source, structure, and/or cell size. Panuska et al. (1991) and Vieux and Needham (1993) quantified the effects of data structure and cell size on Agricultural Non-Point Source (AGNPS) pollution model inputs, and showed how the computed flow-path lengths and upslope contributing areas varied with element size. Vieux (1993) examined the sensitivity of a direct surface runoff model to the effects of cell size aggregation and smoothing using different-sized windows. Moore et al. (1993d) examined the sensitivity of computed slope and steady-state topographic wetness index values across 22 grid spacings for three moderately large ($\approx$100 km^2) catchments in southeastern Australia. Hodgson (1995) demonstrated that

the slopes and aspects calculated from 30-m DEMs are representative of grid spacings two or three times larger than the original DEM grid spacing. Issacson and Ripple (1991) compared 1° USGS 3-arc-second and 7.5′ USGS 30-m DEMs and Lagacherie et al. (1996) examined the effect of DEM data source and sampling pattern on computed topographic attributes and the performance of a terrain-based hydrology model. Chairat and Delleur (1993) quantified the effects of DEM resolution and contour length on the distribution of the topographic wetness index as used by TOPMODEL and the model's peak flow predictions. Wolock and Price (1994) and Zhang and Montgomery (1994) also examined the effects of DEM source scale and DEM cell spacing on the topographic wetness index and TOPMODEL watershed model predictions. Garbrecht and Martz (1994) examined the impact of DEM resolution on extracted drainage properties for an 84-km^2 study area in Oklahoma using hypothetical drainage network configurations and DEMs of increasing size. They derived various quantitative relationships and concluded that the grid spacing must be selected relative to the size of the smallest drainage features that are considered important for the work at hand. Bates et al. (1998) showed how high-frequency information is lost at progressively larger grid spacings.

1.2 THE PURPOSE OF THIS BOOK

The preceding review is instructive in at least four ways. First, it highlights the tremendous interest in digital terrain analysis that has emerged during the past decade and the key contributions of Ian D. Moore during the period 1985–1993. Second, it describes the most popular topographic attributes and the methods that have been used to calculate them. Third, it illustrates some of the ways in which the computed topographic attributes have been used to improve our understanding of hydrological, geomorphological, and ecological systems. Finally, it describes many of the subtleties and challenges that must be overcome in order to use digital elevation data and terrain-analysis tools effectively.

Most of the chapters in this book look past the problems raised at the end of the previous section and demonstrate some of the ways in which these continuously varying but gridded landform attributes can be used to quantify topographic controls on hydrological, geomorphological, and ecological systems. The individual chapters included in this book describe the TAPES terrain-analysis methods and show how computed terrain attributes can be utilized to describe key environmental patterns as a function of process. The applications confirm that the current methods and data sources are best suited to work at intermediate spatial (hillslopes and catchments) and temporal scales (measured in terms of months or years).

In all of this work, we must take care to ensure that our simplifying assumptions resolve rather than introduce computational complexity. Figure 1.4 is adapted from a similar diagram in Grayson et al. (1993) and shows how the terrain-based "hydrological information content" can be expected to vary with changing element size (i.e., spatial resolution). The relationship in the right-hand side of Figure 1.4 shows how the quantity of information declines as the element size increases beyond the scale of the

measurements (topographic attributes in this instance). This is a generally accepted notion and is caused by the lumping of subgrid information (I. D. Moore et al. 1991, Bates et al. 1998). Two scenarios are captured as element size is reduced on the left-hand side of Figure 1.4. In one instance, the topography controls the lateral migration and accumulation of water and the finer resolution increases the information content. In the second case, the hydrological behavior is dominated by soil characteristics, such as preferential flow paths, that are not related to topography and the finer spatial resolution does not increase the level of hydrological information (Grayson et al. 1993). These examples suggest at least three challenges. One is the need to increase our understanding of the key processes affecting sediment and water behavior at a variety of scales (e.g., Moore and Grayson 1991, Grayson and Moore 1991, Grayson et al. 1992a, Robinson et al. 1995). The second is concerned with the development and testing of methods to measure and/or interpolate values for these variables across landscapes (Phillips 1988). The final challenge involves the identification of indicators that can be used to monitor changes in individual processes (Kirkby et al. 1996).

The individual chapters included in this book raise many of the conceptual and methodological issues that we will need to consider as we forge ahead with these

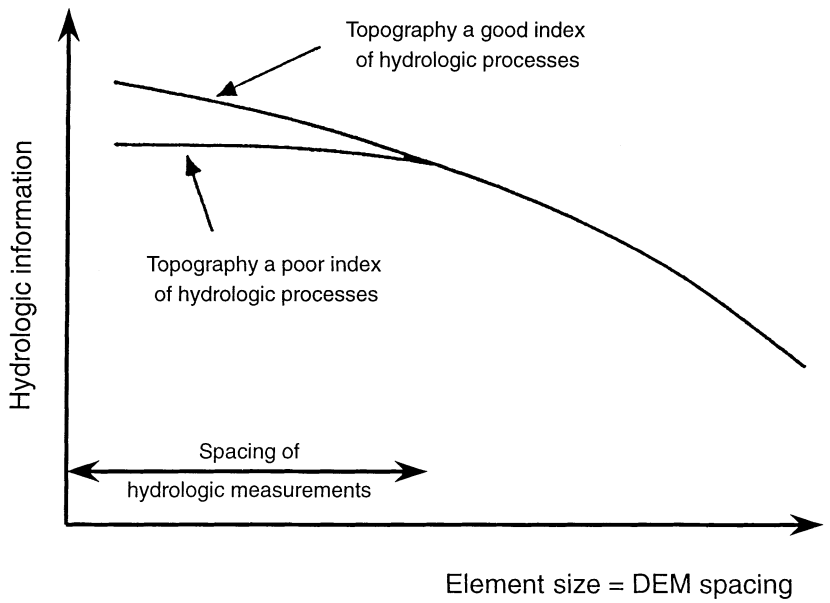

Figure 1.4. Conceptual representation of the relationship between hydrologic information and element size for topography-based interpolations. Reprinted with permission from Grayson, Blöschl, Barling, and Moore (1993) Process, scale, and constraints to hydrological modeling in GIS. p. 83–92 in *Application of Geographic Information Systems in Hydrology and Water Resources: Proceedings of the HydroGIS '93 Conference held in Vienna, April 1993*, edited by K Kovar and H P Nachtnebel. Copyright © 1993 by International Association of Hydrological Sciences, Wallingford, United Kingdom.

types of applications in the future. The applications presented in this book also demonstrate how simple spatial models can be combined with qualitative reasoning to improve our understanding and management of environmental systems. Grayson et al. (1993) and others have argued that this approach is consistent with our current ability to represent biophysical systems and the availability of data. We hope that the methods and applications described in this book will encourage others to adopt this approach and thereby help to increase the quantity and quality of the scientific concepts and geospatial information used in environmental assessments and management applications.

1.3 OVERVIEW

This book is divided into four main sections. The first section is concerned with methods and data and the remaining three sections illustrate hydrological, geomorphological, and biological applications, respectively. The final chapter offers some concluding remarks and future predictions. Our contributions consist of two chapters in addition to this one and co-authorship of five other chapters. The remainder of the authors were selected because of their utilization of the TAPES software tools and participation in a workshop celebrating the life and scientific contributions of Ian D. Moore (1951–1993) that we organized at the Third International Conference Integrating GIS and Environmental Modeling in 1996. Their contributions have evolved in various ways in the three years that have elapsed since the workshop and the following subsections summarize the contents of the individual chapters.

1.3.1 Digital Terrain Analysis Methods

This initial chapter and the four that follow describe the conceptual foundations and methods of digital terrain analysis. These chapters therefore provide important background material for the subsequent chapters on hydrological, geomorphological, and biological applications in this book, and for readers who may be interested in similar types of applications. Table 1.3 lists the individual terrain-analysis programs by book chapter for those interested in specific topographic attributes and/or terrain-analysis methods.

Michael Hutchinson and John Gallant describe DEM data sources and interpolation methods in relation to the accurate representation of terrain shape and scales of source data and applications in Chapter 2. A contour map for the Cottonwood Creek catchment on the Red Bluff Montana Agricultural Experiment Station was digitized and is used to illustrate the application of the interpolation methods described in this chapter. The contour map and a 15-m square-grid DEM that was produced from it are used in Chapters 3 and 4 to illustrate the application of six additional terrain-analysis programs.

We describe the methods used by TAPES-C and TAPES-G to calculate a series of primary topographic attributes in Chapter 3. TAPES-C and TAPES-G start with contours and regular grids, respectively. Both programs generate spatially variable esti-

TABLE 1.3 Utilization of Individual Terrain Analysis Programs by Book Chapter

Chapter	TAPES-C	TAPES-G	DYNWET	EROS	SRAD	WET
3	X	X				
4			X	X	X	X
5		X		X		
6		X				X
7		X	X			
8	X					
9	X					
10		X				X
11		X				
12		X				
13		X				
14		X				
15		X	X		X	
16		X			X	

mates of slope, aspect, profile and plan curvature, flow-path length, specific catchment area, and several other topographic attributes. The inputs, estimation methods, and outputs are described for each program in turn, and the 15-m Cottonwood Creek DEM produced in Chapter 2 is used to illustrate the performance of both programs.

We describe the methods used by four grid-based programs (EROS, SRAD, WET, DYNWET) to calculate several sets of secondary topographic attributes in Chapter 4. EROS estimates the spatial distribution of soil loss and erosion and deposition potential in a catchment. SRAD computes the radiation budget using incoming short-wave irradiance and incoming and outgoing long-wave irradiance for periods ranging from 1 day to a year. Slope, aspect, topographic shading, and monthly variations in cloudiness, atmospheric transmissivity, and vegetation properties are taken into account, and this program can also be used to estimate surface and minimum, maximum, and average air temperatures. WET calculates equilibrium soil moisture, evaporation, and runoff based on topographic attributes and spatial estimates of solar radiation, and DYNWET calculates topographic wetness indices based on both the steady-state and quasi-dynamic subsurface flow assumptions. The inputs, estimation methods, and outputs are described for each program in turn, and the 15-m Cottonwood Creek DEM produced in Chapter 2 is used to illustrate the performance of all four programs.

John Wilson, Philip Repetto, and Robert Snyder examined the effect of DEM data source, grid resolution, and flowing routing method on computed topographic attributes in a pair of experiments conducted in southwest Montana and northern Idaho (Chapter 5). The Montana experiment quantified the sensitivity of five attributes in the 105-km^2 Squaw Creek catchment. The agreement between maps of computed topographic attributes derived from DEMs of different size was poor when the attributes were reclassified into five classes (28–49% agreement) and slightly better when 7.5′ USGS 30-m DEMs were used with different flow-routing algorithms (49–71%

agreement). The Idaho experiment was conducted on a single farm field and examined topographic controls on soil erosion and the ability of the stream-power index used with 7.5′ USGS 30-m DEMs to distinguish field areas experiencing net erosion and net deposition. Numerous explanations are offered for the poor performance of the stream-power index in this particular instance. Overall, the results from these experiments showed why care must be exercised when choosing DEM sources, grid resolutions, and terrain-analysis methods for different types of applications and landscapes, and in that sense, they corroborate many of the arguments raised in Chapters 1 and 2.

1.3.2 Hydrological Applications

The four chapters in this section incorporate one or more topographic attributes in a variety of GIS-based hydrologic modeling frameworks. All four applications explore the topographic influences on soil moisture and runoff behavior, but each uses the computed topographic attributes with successively more complicated hydrologic models and smaller study areas. Hence, the four study areas used in these applications ranged from 96,000 km^2 to 2.5 ha in spatial extent.

In Chapter 6 Valentina Krysanova, Dirk-Ingmar Müller-Wohlfeil, Wolfgang Cramer, and Alfred Becker combined the WET model and a variant of the USLE with a GIS to explore spatial patterns of soil moisture and potential soil loss in the 96,000-km^2 German portion of the Elbe River Basin. The models are used to identify the areas that are expected to experience long-term average soil moisture deficits and accelerated soil erosion as a function of long-term average climate, topography, soil, and land-use data. The maps of wetness index generated with WET showed good agreement with maps of long-term water availability expressed as groundwater table depths and the map of soil loss potential showed good agreement with previous studies. This application illustrates how the computed topographic attributes can be used to identify areas that warrant more detailed analysis.

Jeremy Fried, Daniel Brown, Mark Zweifler, and Michael Gold (Chapter 7) construct a series of stream buffers using cumulative cost distance calculated over fuzzy set combinations of relative topographic wetness and stream-power indices for a 17-km^2 first-order Michigan drainage basin. Their investigative buffers (models) are grounded on the assumption that riparian segments receiving the greatest discharge have upslope contributing areas dominated by saturated soils and have sufficient stream power for saturated flow to reach the stream. The resultant models are evaluated using field data collected during a post-storm-event GPS field survey of ponded storm flow accumulations and concentrated storm flow discharge sites. This application is instructive because it shows how topographic indices might be combined with qualitative reasoning to guide site-specific water pollution remediation efforts in the future.

Alan Yeakley, George Hornberger, Wayne Swank, Paul Bolstad, and James Vose describe the development, calibration, and testing of a terrain-based hillslope hydrology modeling framework for simulating soil moisture distributions in forested landscapes in the southern Appalachian Mountains (Chapter 8). Their approach

incorporates an above-ground interception model, contour-based topographic attributes computed with TAPES-C, and a watershed hydrology model. These models were applied to a 12.3-ha first-order drainage basin that is part of the Coweeta Hydrologic Laboratory, where previous work has demonstrated the dominance of subsurface flow and absence of overland flow due to the high infiltration capacities of the forest soils. The new modeling framework captured the mean soil moisture response during storm events and over several years (seasons) but did less well at capturing soil moisture extremes and depicting spatial variability. These shortcomings indicate why increased knowledge of the spatial variation in soil hydraulic properties and more accurate representation of soil moisture near seepage faces are needed to build more successful (complete) explanatory models.

In the final chapter in this set (Chapter 9), Greg Pohll and John Warwick describe some work to build an improved explanatory model for the hydrologic behavior of a subsidence crater at the Nevada Nuclear Test Site. An increased understanding of the linkages between surface and subsurface components is required to predict large-scale radionuclide transport and to assess the feasibility of using subsidence craters as low-level nuclear waste storage sites. The crater used in this particular study has a diameter of 180 m and drains a 2.5-ha catchment. The TAPES-C terrain analysis program was dynamically linked with an overland flow simulator (THALES) and a numerical Richards' equation solver (SWMS-2D) to represent moisture migration within the vadose zone and simulate the movement of water during overland flow, ponding, infiltration, and seepage at the study site. The outputs were compared with a simpler vadose zone model with static surface boundary conditions to determine the effectiveness of each model. The results were mixed—the simpler and more complicated models offered superior predictions of deep moisture migration and the temporal distribution of infiltration, respectively—and, like the results from the previous chapter, demonstrate the difficulty of building realistic hydrologic process models.

1.3.3 Geomorphological Applications

The four chapters in this section examine pedological and geomorphological applications of digital terrain analysis. The breadth of the applications declines from one chapter to the next. The first chapter starts with a review of theories of pedogenesis and the role of soil survey in summarizing and communicating knowledge of soil properties, and the last chapter describes a terrain-based model to delineate shallow landslide areas in steep forested watersheds. The study areas used in the four chapters vary tremendously in size and character as well.

In the first chapter in this set (Chapter 10), Neil McKenzie, Paul Gessler, Philip Ryan, and Deborah O'Connell review the role of soil survey and describe a series of soil survey applications from southeastern Australia. Digital terrain analysis is used to characterize microclimates, develop explicit statistical sampling plans, and generate spatial predictions of soil properties at resolutions unmatched by comparable conventional methods for the 500-km^2 Bago-Maragle study area in southern New South Wales. These authors argue that digital terrain analysis has created an opportunity for a more scientifically based method of soil survey and that they may (one

day) be used to generate improved quantitative spatial predictions of specific soil properties. The next two chapters illustrate both of these possibilities in specific environments.

Stephen Ventura and Barbara Irvin (Chapter 11) describe the use of topographic attributes to automate the classification of landform elements for a 49-ha study area in the "Driftless Area" of Wisconsin. Continuous (fuzzy logic) and unsupervised classification techniques were used to assign DEM cells a membership of a landform element class. The unsupervised classification assigned cells to single landform classes, whereas the continuous classification allocated relative class memberships for every class in every cell. These classes were determined by the natural clustering of the data in attribute space in both instances. The clusters formed readily recognizable patches on the landscape that matched manually interpreted landform classes and soil survey map units.

Jay Bell, David Grigal, and Peter Bates (Chapter 12) developed several quantitative models that predict soil organic carbon (SOC) storage as a function of topographic attributes and vegetative cover for a 22-km^2 study area in the Cedar Creek Natural History Area of Minnesota. Different models were constructed for mineral soils and peatlands, and the final model described approximately 50% of the variation in SOC over the entire study area. Slope and several relative elevation and distance measures were included as explanatory variables in the final models.

Jinfan Duan and Gordon Grant used topographic attributes with an infinite slope model to predict shallow landform areas in the final chapter in this set (Chapter 13). Their approach incorporated a dynamic simulation of rainfall intensities and treated the spatial distribution of key soil and vegetation parameters stochastically using a Monte Carlo simulation approach. This model was tested using observed landslides for a 64-km^2 drainage basin in western Oregon. The agreement with the locations of observed slides was only fair, and the authors concluded that their model may be most useful in predicting average slide frequencies under different management regimes rather than identifying specific slide locations, as other, more deterministic models do. This application is similar to the final two hydrologic applications in Chapters 8 and 9 in that it highlights some of the difficulties that are encountered in using static topographic indices to represent dynamic landscape processes.

1.3.4 Biological Applications

The three chapters in this set examine the linkages between computed topographic attributes and vegetation patterns in three North American landscapes.

Janet Franklin, Paul McCullogh, and Curtis Gray developed classification tree models relating topographic attributes and spectral variables derived from satellite imagery to chaparral species associations and riparian vegetation types for a study area in the Laguna Mountains of the Peninsula Ranges in San Diego County, California (Chapter 14). The models were then applied to digital maps of the topographic and spectral variables to produce predictive maps of vegetation distribution with accuracy estimated in the 52–62% range. Very detailed vegetation types (classes) were identified, and the authors concluded that their approach provides a classifica-

tion method that avoids the need for a priori judgments about the terrain/satellite imagery/vegetation relationships.

Jonathan Wheatley, John Wilson, Roland Redmond, Zhenkui Ma, and Jeff DiBenedetto examined whether one or more topographic attributes can be added to an existing satellite interpretation method to improve land cover classification accuracy in the Little Missouri Grassland of North Dakota (Chapter 15). Quasi-dynamic topographic wetness and incident short-wave solar radiation indices were added to the third stage of a four-step *Landsat* TM satellite interpretation method. Error matrices were developed using a "bootstrap" process that removed each plot from the training data set, one at a time, and used the remaining plots to classify each one. Accuracy remained in the 51–57% range using 13 land cover types and 173 ground-truth plots. The terrain-analysis tools did help with the identification of the channel system and the results confirmed why satellite-based land cover maps must be used with care, since different source data and levels of spatial aggregation will predict different patterns of existing vegetation.

Brendan Mackey, Ian Mullen, Kenneth Baldwin, John Gallant, Richard Sims, and Daniel McKenney examined topographic controls of boreal forest ecosystems in the Rinker Lake region of northwestern Ontario, Canada (Chapter 16). A nonparametric statistical model is used to correlate the distribution of Jack Pine with environmental field measurements and computed topographic attributes. The results showed that topographic indices derived from a 20-m DEM were better predictors of Jack Pine than in situ observations of either substrate or topography. This implies that there are strong topographic controls on the distribution of Jack Pine in the Rinker Lake area. Various explanations are offered for this relationship and they all point to the need for additional information and knowledge if topographic attributes are to be interpreted and used correctly.