9

Approaches to spatially distributed hydrological modelling in a GIS environment

L. Olsson and P. Pilesjö[1]

ABSTRACT

Traditionally hydrological models have been based on the drainage basin as the fundamental system delineation and their function have been empirically based. In order to build models capable of describing the movement of water within a drainage basin, a spatially explicit approach is needed (see Chapter 2 for an overview of model types). This paper deals with two aspects of the spatially distributed hydrological model – the atmospheric interface and the geomorphological distribution of water flow. Recent advances in land surface-atmosphere interaction models have improved substantially our ability to estimate the fluxes of water and energy between the atmosphere and the terrestrial ecosystems. One of the most important developments is the use of remotely sensed data for measuring, in a spatially continuous fashion, land surface parameters that can serve as input data to models. Coupling of SVAT (Soil Vegetation Atmosphere Transfer) models with distributed hydrological process (see Chapter 2) models and biological production models can be used to assess the effects of land use/cover changes on the regional hydrological cycle. Recent developments in the processing of digital elevation models for the estimation of flow accumulation and automatic delineation of drainage basins is another important basis for the development of spatially distributed models.

9.1 BASIC HYDROLOGICAL PROCESSES AND MODELLING APPROACHES

Water is the most important limiting factor to vegetation growth, and thereby also one of the most important factors controlling human livelihood. Consequently, modelling within the hydrological cycle has become one of the most important tasks in terrestrial ecology. The aim of the traditional hydrological model has primarily been to predict the amount of discharge from a drainage basin, while water movement within the basin has often been neglected. With the advent of efficient computers and spatial data of high quality, the interest has shifted from those lumped models (see Chapter 2) towards spatially distributed models, where

[1] Note that the authors contributed equally to the paper.

water movement within the drainage basin can be modelled. Important applications of this emerging field of spatially distributed hydrological modelling tools include studies of:

- pollution propagation in the soil
- impact of land surface (e.g. agriculture and forestry) management practices on hydrological regimes
- impact of vegetation and land use change on hydrological regimes
- the prediction of nutrient leakage in agricultural landscapes.

The aim of this chapter is to outline the design of distributed models in a GIS environment and to discuss problems and potentials. Another aim is to discuss the use of remote sensing data as an aid in modelling hydrological processes.

The development of a spatially distributed hydrological model can be described as solving three major problems, to be solved in a geographically explicit fashion, these are:

- the partitioning of precipitation into evaporation and water input to the drainage basin
- the partitioning of water input into infiltration and surface runoff
- the movement of surface and subsurface water within the drainage basin.

In the first part of the chapter we describe generally the different components of the hydrological cycle. Understanding of the fundamental processes of water flow is essential in all types of modelling. Even if all models are only generalized mimics of the environment, profound knowledge about the processes helps us to develop and evaluate the models. Apart from the processes, the introductory part also describes different modelling approaches. This section is based on Andersson and Nilsson (1998).

In the second part of the chapter the main input parameters to hydrological models, and mechanisms by which these are derived using GIS and remote sensing, are described. The following section deals with the land surface – atmosphere interface, which describes different model approaches to divide the precipitation into evapotranspiration and water input to the drainage basin. The partitioning of water into infiltration and surface runoff as well as the movement of surface and subsurface water is discussed at the end of the chapter.

9.1.1 The hydrological cycle

There is an unending circulation of water within the environment. This circulation is called the hydrological cycle and its components are presented in Figure 9.1. The energy from the sun evaporates water from open water surfaces and from land. Wind transports the moist air until it condenses into clouds in a cooler environment. Water reaches the ground and water surfaces as precipitation (rain, snow or hail) falling from the clouds. Depending on temperature, the precipitation can be stored as snow, ice or water for a shorter or longer time. A part of the falling rain is captured by the vegetation as intercepted water, which eventually evaporates

back to the air. The remainder of the rain will reach the ground or fall into water bodies where evaporation will continue. A portion of the water that reaches the ground may flow directly into streams as overland flow, but most of the surface water will infiltrate into the soil. The water infiltrates until the soil is saturated and cannot hold any more water. If the soil becomes saturated, the excess water will flow on the soil surface as overland flow. The water that has infiltrated into the soil will move downwards or laterally as subsurface flow. The lateral movement is due to diversion when soils of different characteristics are reached or because of differences in water pressure in the water saturated zone (ground water). The downward movement is due to gravity and the water will eventually become part of the ground water. The infiltrating water may also be taken up by vegetation from which it may be transpired back to the atmosphere.

Subsurface, ground water and overland flows contribute to the stream flow which transports the water back to the ocean and completes the hydrological cycle.

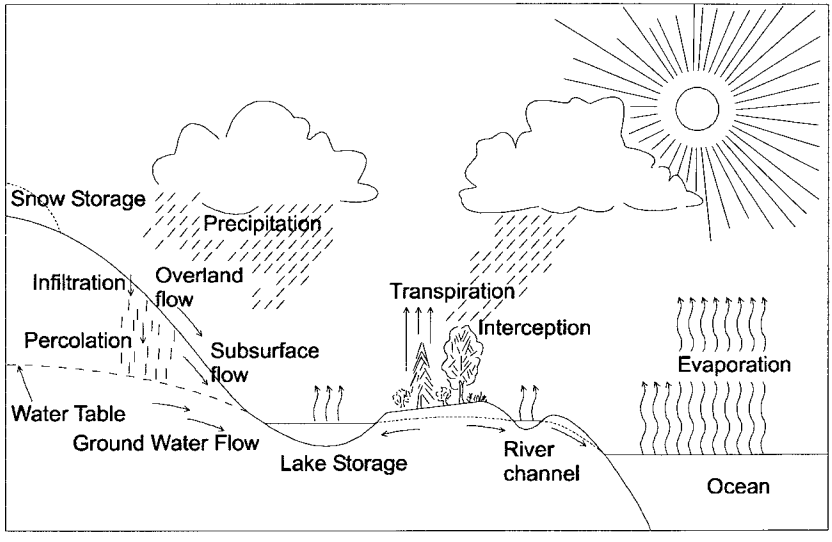


Figure 9.1: Components of the hydrological cycle (from Andersson and Nilsson 1998).

Below follows a short description of the most important components of the hydrological cycle:

Precipitation occurs when moist air is cooled and reaches the dew point temperature. This gives rise to water droplet development on condensation nuclei, e.g. small dust particles.

Precipitation intensity is the amount of precipitation (in liquid form) per time unit.

Interception occurs when vegetation captures precipitation on its path to the ground. The capacity to store water on leaves and stems depends on vegetation type and appearance. The capacity is generally higher for evergreens than for deciduous trees (Selby 1982).

Evaporation is used as a term for the loss of water vapour from water, soil and vegetation surfaces to the atmosphere. The process deals with changes in state of aggregation from liquid water into water vapour and is controlled by the moisture gradient between the surface and the surrounding air. The energy used for the change is primarily net radiation, i.e. from the sun, but can also be taken from stored heat in e.g. vegetation or water bodies.

The water vapour capacity of the air is directly related to temperature (Shaw 1993).

Evaporation from soil originates from temporary surface puddles or from soil layers near the surface. The effectiveness of the evaporation depends on the aerodynamic resistance, which in turn is dependent upon wind speed, surface roughness, and atmospheric stability, all of which contribute to the level of wind turbulence (Oke 1995).

Transpiration is the water loss from the soil through the vegetation. Transpiration differs from evaporation since vegetation can control its loss of water. Plants draw their water supply from the soil, where the moisture is held under pressure. They control the rate of transpiration through the stomata in their leaves by changing the area of pore openings. Usually this factor is referred to as stomata resistance, and depends on the water content of the air, the ambient temperature, the water availability at root level, light conditions, and carbon dioxide concentration (Oke 1995). The pores close in darkness and hence transpiration ceases at night. When there is a shortage of water in the soil the stomata regulates the pores and reduces transpiration. Transpiration is thus controlled by soil moisture content and the capacity of the plant to transpire, which in turn are conditioned by meteorological factors (Shaw 1993).

If there is a continuous supply and the rate of evaporation is unaffected by lack of water, then both evaporation and transpiration are regulated by the meteorological variables radiation, temperature, vapour pressure and wind speed (Shaw 1993). When the vegetation is wet the loss of water is dominantly due to evaporation. During dry conditions the water loss from vegetation surfaces is mainly via transpiration. Some 20–30 per cent of the evaporated water originates from intercepted vegetation storage when such occur (Lindström *et al.* 1996).

Usually the combined loss of water from ground, water surfaces and vegetation to the atmosphere is called *evapotranspiration.*

Infiltration is generally described as the penetration and flow of water into the soil. When a soil is below field capacity, which is the capacity of water content of the soil after the saturated soil has drained under gravity to equilibrium, and precipitation is gathered on the surface, the water penetrates into the soil. The water infiltrates at an initial rate dependent on the actual soil moisture content and the texture and structure of the soil. As the precipitation supply continues the rate of infiltration decreases, as the soil becomes wetter and less able to take up water. The typical curve of infiltration rate with time reduces to a constant value, called the infiltration capacity (Shaw 1993), which usually is equal to, or slightly less than, the saturated hydraulic conductivity. The hydraulic conductivity is a measure of the water leading capability of the soil, and is controlled by the soil pore size, soil composition, and the soil moisture content. The saturated hydraulic conductivity is often referred to as permeability (Grip and Rhode 1994).

The actual infiltration capacity of a soil varies depending on the soil characteristics and the soil moisture. Pre-existing soil moisture is an important infiltration regulating factor because some soils exhibit an initial resistance to wetting (Selby 1982).

During infiltration the soil is getting soaked, but when the rainfall stops soil beneath the wetting front is still getting wetter while soil above is drying as it drains. The type of vegetation cover is also an important factor influencing infiltration. Denser vegetation results in higher organic concentration in the root zone, promoting a thicker soil cover and a more loose structure. This results in a higher infiltration capacity. Vegetation and litter also decrease the precipitation impact on the surface. Without these factors smaller particles would be thrown into suspension, and clogging might occur as they are re-deposited and less permeable layers would evolve (Chorley 1977).

Overland flow is often divided into Hortonian overland flow and saturated overland flow. When precipitation intensity exceeds the infiltration capacity, the precipitation still falling on the area can not infiltrate and the excess water flows on the surface. This kind of flow is referred to as Hortonian overland flow. The second type of overland flow occurs when completely saturated soils give rise to saturated overland flow without having precipitation falling upon it, due to pressure effects induced by subsurface water infiltrated in up-slope areas. The velocity of overland flow depends on the slope angle and the surface roughness.

Subsurface flow occurs below the soil surface. Close to the surface, where the soil normally is not saturated, we have an unsaturated water flow. At greater depths the soil reaches saturation (i.e. the water pressure exceeds the atmospheric pressure). The surface at which the pressure equals the atmospheric pressure is defined as the groundwater table (see Figure 9.2). Below the groundwater table all soil pores are completely filled with water. This is referred to as the saturated zone. The connected pore system in a soil can be seen as small pipe shaped areas and the water level rise is referred to as the capillary rise. The zone is often called the capillary fringe (Grip and Rhode 1994). The extent of the capillary fringe is dependent on the soil composition and the packing of the soil particles. It ranges from a few centimetres in a coarse sandy soil to several meters in a clay soil. The soil above the capillary fringe is referred to as the unsaturated zone or the aeration zone and has pores filled with a mixture of water, water vapour and air. After e.g. heavy rains, parts of the unsaturated zone might become temporarily saturated.

As shown in Figure 9.2 unsaturated flow occurs in the unsaturated zone. The water with a vertical flow inside the unsaturated soil is usually referred to as *percolation*. The velocity of the diverted flow is dependent on soil permeability and stratum slope (Brady and Weil 1996).

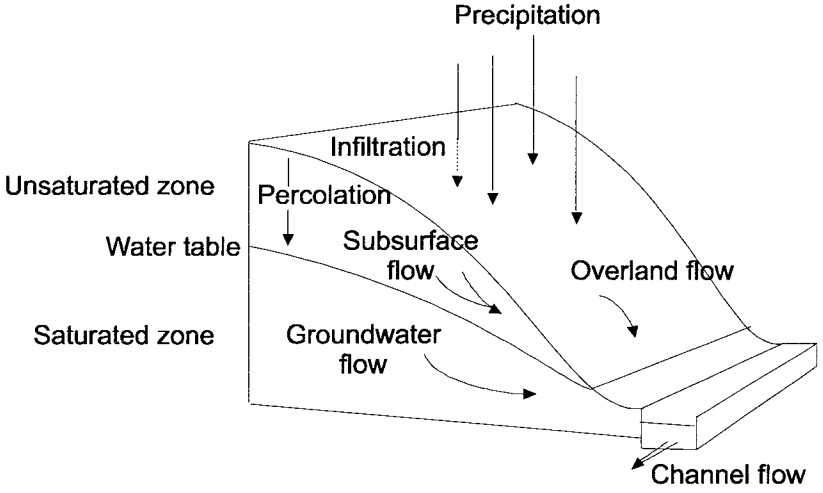


Figure 9.2: Overland and subsurface flow (from Andersson and Nilsson 1998).

In an unstratified soil there is also a tendency for more compaction and smaller pores with greater depth. This leads to successively lesser permeability and a greater partition of the water will be forced to move sideways as subsurface flow (Chorley 1977). The diverted flow is saturated, but not included in the saturated zone since there is unsaturated soil beneath it (Andersson and Burt 1985).

When the percolating water reaches the saturated zone, or more exactly the capillary fringe, the water is incorporated with the groundwater (see Figure 9.2) and the water table is temporarily raised. Inside the saturated zone spatially different water pressures govern the water movements. The theory for groundwater movement is based on Darcy's law (see below). An important extension of Darcy's law for groundwater flow is its application in three dimensions. The permeable material, the soil, is often heterogeneous. Clay layers can for example be present in sandy material, and soil close to the surface is often more porous than material at greater depth. The hydraulic properties of the ground are not isotropic, and the hydraulic conductivity is different in different directions (Bengtsson 1997). The discharge rate is therefore into three perpendicular discharges (Q_x, Q_y and Q_z) with different hydraulic conductivity and different flow velocities (Shaw 1993).

The groundwater flow is always in motion but at very slow velocities, about three magnitudes lower than overland flow (Chorley 1977). At a large scale the groundwater movement is directed from a recharge area to a discharge area. Recharge areas can be defined as areas with a vertical flow component downwards inside the groundwater zone. Discharge areas are defined as the opposite phenomena, where the principal water flow is upwards.

Once water has infiltrated into the ground, its downward movement to the groundwater and the amount of stored groundwater depends on the geological structure as well as on the rock composition. In general, older rock formations are more consolidated and the rock material is less likely to contain water. Igneous and metamorphic rocks are not good sources of groundwater, unless weathered and fractured. The sedimentary rock strata have different composition and porosity and are much more likely to contain large amounts of water (Shaw 1993). Beds of rock

with high porosity that are capable of holding large quantities of water are often referred to as aquifers. Aquitards are semi-porous beds, which allow some seepage of water through them. Clay beds, which are almost impermeable, are called aquicludes (Shaw 1993).

9.1.2 Modelling approaches

The two classical types of hydrological models are the deterministic and the stochastic, where stochastic models involve random elements (see also Chapter 2). The deterministic models can be classified according to whether the model gives a lumped or a distributed description of the considered study area. The models can also be classified whether the description of the hydrological processes is empirical or physically-based. There are three major types of deterministic models: empirical lumped models, empirical distributed models, and physically-based distributed models. The fourth combinatory possible type would be the physical lumped model, but this concept is somewhat contradictory since physical models require measurable input data whereas lumped models use averages for an entire catchment. Classifications of this kind are however fluent. Lumped models might have parameters that are more or less distributed. Models can also have components with both physical and empirical origins, so called semi-empirical or grey box models (Abbott and Refsgaard 1996).

Empirical models (see section 2.4.1) are based on regression and correlation results from statistical analyses of time series data. The derived equations are based on observed phenomena or measurement knowledge without demands on understanding of the underlying processes. Empirical models are often referred to as black box models. Truly *physical models* (see section 2.4.3) are based on formulas of physical relations. They are analogously referred to as white box models (Kirby *et al.* 1993) since every part of the processes is understood. The input data include only measurable variables that can be collaborated.

Physically-based models are the most suitable when studying internal catchment change scenarios. Examples of this are irrigation and groundwater use development. The prediction of discharge from catchments including monitoring of pollutants and sediments dispersed by water are also well suited for physical models (Andersson and Burt 1985; Abbot and Refsgaard 1996). It is important to note that not all of the conceptual understanding of the way hydrological systems work is expressible in formal mathematical terms. Thus any model definition will be an abstraction of the total knowledge of catchment hydrology. Thereby all models include a systematic error based on the not included or not known relationship. This is a neglected source of error in many physical modelling processes, and yields a need of calibrating the model to time series data. Practically, this means we have very few truly physically-based models but many semi-physical ones.

A *lumped model* (see section (2.4.3) operates with interrelated reservoirs representing physical elements in a catchment being the smallest spatial element in the modelling system. This results in that the model uses parameters and variables that represent average values for the entire catchment. These averages can be

derived either physically or empirically which can give the model a semi-empirical appearance. Lumped models are mainly used in rainfall-runoff modelling.

Distributed hydrological models (see section (2.4.3) are supposed to describe flow processes in each and every point inside a catchment. Due to difficulties within the general conceptual modelling framework and very time and memory consuming programs these models are practically impossible to use. Simpler models instead try to estimate the different flow patterns discretisised into nodes with orthographic spacing. These nodes can be seen as centre points in square shaped areas referred to as pixels or cells. If a model is based on this type of cell structure it is directly compatible with remotely sensed and gridded (raster) GIS data. In the vertical extent each orthographic cell might be given a depth, or be discretisised into a number of overlaying cells (i.e. a column). For each cell the water discharge to neighbouring cells is calculated according to the active hydrological processes. The flow distribution inside the catchment is thereby mapped. Even if the processes are estimated as a continuum, the stored results are discretisised into cells (Abbott and Refsgaard 1996).

The distributed nature of a modelling system means that spatial variation, characteristics and changes can be simulated and estimated inside a catchment. Distributed hydrological models have particular advantages in the study of the effects of land use changes. The model not only provides a single outlet discharge, but multiple outputs on a temporally and spatially distributed basis. The disadvantages with this form of modelling are the large amounts of data and the heavy computational requirements. The model type also includes a large number of parameters and variables, which have to be evaluated. The effect of scale choice (cell size) is also an uncertainty (Beven and Moore 1993).

A *stochastic model* (see section 2.5) uses random elements, which are drawn from statistically possible distributions. This means that the simulations will not give the same results when repeated with the same input data. With most stochastic models the approach is to conduct a multitude of simulations, the so-called Monte Carlo technique, and produce average estimates with specified confidence intervals.

9.2 DATA FOR SPATIALLY DISTRIBUTED HYDROLOGICAL MODELLING

One of the most severe problems to overcome in distributed hydrological modelling is the mismatch of scales between processes and obtainable data, both in terms of spatial scales as well as temporal scales. The most important divide in relation to data sources is between point data and spatially continuous data. Most of the climatic data necessary for hydrological modelling can only be obtained at a point basis, even though remote sensing methods are becoming increasingly important. But on the other hand, the point data are often available at very short time intervals (hours). Data on subsurface properties, soil and rock conditions, are also primarily point based and subsequently extrapolated to cover a region. Concerning vegetation, topography and surface conditions, spatial continuously data are often available, but with varying resolution in time and space. Data on topography necessary for the studies of water movement within a catchment are typically available at a spatial resolution of 25 m to 50 m, which corresponds well with data

on vegetation and surface conditions available from high-resolution remote sensing (e.g. Landsat and Spot). However, the temporal resolution of these remote sensing data (typically yearly, considering costs and other practical factors) are too coarse to capture the biological aspects of the hydrological cycle. Temporal resolution adequate for studying vegetation and climatic processes (from bihourly to bimonthly) are available, but at a much coarser resolution, typically 1 to 5 km. In order to make the best out of these conflicting scales, in time and space, profound knowledge on data sources and handling coupled with a large portion of creativity are needed.

9.2.1 Vegetation

In order to successfully set up and run a distributed hydrological model the vegetation must be described by appropriate parameters in a spatially explicit fashion. Over large regions, the only practical means is by remote sensing. The vegetation parameters needed relate primarily to the role of vegetation in the following processes:

- evaporation and transpiration
- interception
- infiltration .

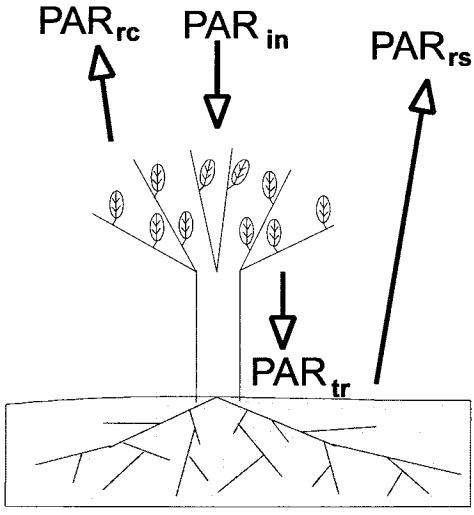
Here we will concentrate on the first two processes.

We can distinguish between two approaches to estimate vegetation parameters from remote sensing (see also Chapter 6):

- to infer vegetation parameters directly from the remote sensing data, or
- to classify vegetation types and model vegetation parameters independently of the remote sensing data.

The first approach requires time series of remote sensing data throughout the vegetation season that is usually only available at a coarse spatial resolution. The most important remote sensing data sources for time series data are the NOAA AVHRR sensing system and different geostationary satellites, e.g. Meteosat covering Europe and Africa. Only the vegetation parameters inferred directly from remote sensing data will be discussed in this chapter.

Satellite sensors provide us with a continuous flow of data on the amount of reflected and emitted radiative energy from the Earth. For studies of vegetation dynamics, the most important part of the spectrum is in the visible and the near infrared region, with special focus on the photosynthetically active radiation (0.4 – 0.7 μm), usually referred to as PAR. Figure 9.3 shows the most important flows of PAR that are used in remote sensing.



PAR_{in} = the incoming amount of photosynthetically active radiation (mW/m^2)
PAR_{tr} = PAR transmitted through the canopy
PAR_{rs} = the amount of PAR that have been reflected by the soil
PAR_{rc} = the amount of PAR that have been reflected by the vegetation canopy

Figure 9.3: The partitioning of incoming photosynthetically active radiation.

We can then define the important parameter absorbed PAR, APAR, according to:

$$APAR = (PAR_{in} + PAR_{rs}) - (PAR_{rc} + PAR_{tr}) \qquad (9.1)$$

Green plants use water (from the roots), carbon dioxide (from the atmosphere) and energy (from the sun) as input to the photosynthesis, and if we can determine the amount of energy the plants are using, we have a link to measure the rate of photosynthesis, which is also an important link to other biophysical processes related to hydrology.

A fundamental problem in remote sensing is to distinguish between the vegetation fraction of the signal from the soil fraction of the signal. This is usually approached through the construction of a vegetation index, where we can distinguish two principally different kinds of indices, ratio-based indices and orthogonal-based indices (for a comprehensive discussion of different vegetation indices, see Huete (1989) and Begue (1993)). The normalized difference vegetation index, NDVI, is the one that has become the most widely used index, and it is defined as equation 9.2.

$$NDVI = \frac{(NIR - RED)}{(NIR + RED)} \qquad (9.2)$$

where *NIR* and *RED* is the amount of reflected light of visible red and near infrared wavelengths respectively.

A number of studies, from ground based measurements as well as from satellite based ones, have confirmed a linear relationship between NDVI and the fraction of absorbed PAR to the incoming PAR, a parameter usually denoted fAPAR and defined as in equation 9.3.

$$fAPAR = \frac{APAR}{PAR_{in}} \qquad (9.3)$$

Some empirical relationships between fAPAR an NDVI presented in the literature are shown below.

fAPAR = 1.42 * NDVI – 0.39 (Lind and Fensholt 1999)
fAPAR = 1.41 * NDVI – 0.40, r2 = 0.963 (Pinter 1992)
fAPAR = 1.67 * NDVI – 0.08 (Prince and Goward 1995)
fAPAR = 1.62 * NDVI – 0.04, r2 = 0.96 (Lind and Fensholt 1999)

Regression slopes found in the literature usually are between 1.2 and 1.6 with intercepts between -0.02 and -0.4. The linear relationship between NDVI and fAPAR have been found to be remarkably consistent over a range of non-woody vegetation types but the relationship deteriorates when LAI becomes more than about 2.

The derivation of leaf area index is less straightforward than the fAPAR transformation above, and the best result is probably obtained by the use of a full radiative transfer model, see for example: Begue (1993) and Myneni and Williams (1994), but for operational use an approach based on Beer's law can be used. Beer's law expresses the relationship between the amount of PAR transmitted through a vegetation canopy and the LAI, according to:

$$PAR_{tr} = e^{\frac{-G \cdot LAI}{\sin b}} \qquad (9.4)$$

where,

b = solar elevation (could be replaced by solar zenith angle, θ (sin b = cos θ))
G = mean direction cosine between solar zenith angle and leaf normals.

A relationship between LAI and fAPAR can then be established (Sellers *et al.* 1996) which scales the LAI logarithmically as a function of fAPAR. The maximum LAI for a particular vegetation type is set to the corresponding maximum fAPAR for that vegetation, according to the formula below.

$$LAI = LAI_{\max} \cdot \frac{\log(1 - fAPAR)}{\log(1 - fAPAR_{\max})} \qquad (9.5)$$

where,

LAI_{max} = the maximum possible LAI for a particular vegetation type
$fAPAR_{max}$ = the corresponding fAPAR for the LAI_{max}.

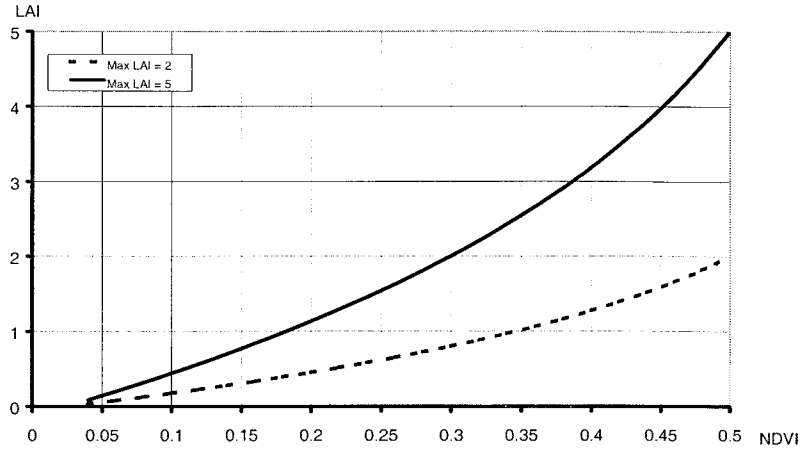


Figure 9.4: Leaf area index as a function of NDVI for two vegetation types with maximum LAI of 2 and 5 and with a corresponding maximum fAPAR of 0.77.

Typical ranges of LAI found using this relationship over the Sahel region in Africa using NOAA AVHRR data are shown in Table 9.1.

Table 9.1: Typical LAI values found in the Sahel region of Northern Africa by transforming NDVI to fAPAR and fAPAR to LAI using equation 9.5. Data from NOAA AVHRR.

Vegetation class	Minimum LAI	Maximum LAI
Evergreen broadleaf forest	0.6	3
Closed shrublands	0.1	0.7
Open shrublands	0	0.15
Woody savannahs	0.4	2
Savannahs	0.2	1.8
Grasslands	0	0.5
Croplands	0	1.7
Cropland and natural vegetation mosaic	0.1	1
Barren or sparsely vegetated	0	0.1

Leaf area index plays an important role in many processes related to how vegetation controls the movement of water from the roots, through the plant and from the plant to the atmosphere. It is also important in determining the amount of water intercepted by a vegetation canopy.

Interception of water by the vegetation canopy is a very complex process to estimate and different approaches to interception estimation have been tried, for a discussion of these, see Gash (1979), Calder (1986) and Jones (1997). One very simple approach has been taken by the SHE model (Abbot and Refsgaard 1996), and is expressed as,

$$I_{max} = C_{int} \cdot LAI \qquad (9.6)$$

where
I_{max} = interception storage capacity
C_{int} = empirical interception parameter, typically 0.05 mm for deciduous forest.

9.2.2 Modelling vegetation growth

It is obvious that a realistic modelling of water balance needs a coupling of water balance models and vegetation growth models. The most promising approach is to use a top-down modelling driven by a combination of remote sensing data and ground observations.

On of the most widely used methods for quantification of vegetation growth is the one that relates net primary productivity (NPP) to the accumulated amount of absorbed photosynthetically active radiation (APAR) over one year or one vegetation season.

$$NPP = \sum_{i=1}^{365} \varepsilon_i \cdot APAR_i \quad (9.7)$$

where
ε_i = a factor accounting for photosynthetic efficiency
$APAR_i$ = absorbed photosynthetic active radiation.

For a remote sensing based solution we may use the previously described relationship between fAPAR and NDVI (equation 9.3).

Combining the equations yields the equation:

$$NPP = \sum_{i=1}^{365.} \left[\varepsilon_i \cdot \left(a \cdot NDVI_i + b \right) \cdot PAR_{in_i} \right] \qquad (9.8)$$

where PAR_{in} = the incoming amount of PAR, which is a function of latitude, time and cloudiness.

The real challenge is then to find ways of estimating the efficiency factor, which is a factor varying with vegetation stress due to water, temperature and nutrient deficiencies. We can distinguish two different approaches to estimate the stress factor, remote sensing based methods involving the use of thermal and visible/NIR sensors (Asrar 1989), and indirect estimation through biophysical models (for example like CENTURY, CASA, SiB, BATS etc.) (Sellers 1992; Prince and Goward 1995; Lind and Fensholt 1999).

9.2.3 Topography

The first step in hydrological modelling is to define a model area by delineating the outline of the catchment boundary. Since our modelling approach is distributed, and more topographic information is needed (e.g. slope and drainage area in order to estimate overland flow, see below) we normally use a digital elevation model (DEM) for the estimations of topographically related parameters.

A DEM is a matrix where every cell value represents the elevation at the centre point in the corresponding area on the Earth's surface. Normally the DEM is interpolated from line or point data (e.g. paper maps or point measurements), but can also be constructed by the use of digital photogrammetry. The quality of the DEM is crucial for the estimation of the topographic parameters as well as for the reliability of the model output.

The quality of the DEM depends on numerous things, but three major factors can be easily distinguished:

- the used interpolation algorithm
- the spatial distribution of the input data points
- the quality (x, y, z) of the input data points.

We can interpolate any spatially distributed variable. After the interpolation, a continuous surface in three dimensions is created. The third dimension of the surface is the value of the measured variable (in our case elevation).

Interpolation algorithms can be divided into two types: global and local methods. The global interpolation methods use all in-data points when estimating the surface, and the results show general trends in the data material. These algorithms are normally not suitable for interpolation of topographical data. Preferably we use a local method, which only use a limited number of in-data points, close to the cell that we are interpolating, when working with topographical data. The obvious reason to use a local method is their sensitivity to smaller, locally distributed terrain forms that are often present in natural terrain. Examples of local interpolation methods are Thiessen polygons, inverse distance interpolation, spline functions and kriging (see Burrough and McDonnell 1998).

The choice of interpolation algorithm primarily depends upon the spatial autocorrelation of the in-data points, and different methods can produce significantly different results. However, the use of geo-statistical interpolation (e.g. kriging), where the user examine the spatial variation in the data and let the autocorrelation in different directions guide the interpolation, is often recommended for interpolation of evenly distributed point data.

As mention above, the distribution of the input data points is extremely important for the result of the interpolation, and consequently for the accuracy of the DEM. Most interpolation algorithms are very sensitive to the distribution of in-data, and demand evenly distributed data to produce a reliable result. Even if some algorithms are more sensitive than others, the importance cannot be stressed enough.

One very useful approach to handle data unreliability is to use Monte Carlo simulation (i.e. a stochastic model – see Chapter 2) when generating a DEM. This is a method which is conceptually very simple, but very computer intensive. The

principle is to repeat a calculation many times, and each time a random element is added to the input data. The method is particularly well suited for applications where it is impossible to analytically calculate the error in every data point, but where it is possible to estimate the overall variability of the data set. One such case is the task to delineate watersheds using DEMs. If we can estimate the overall degree of confidence in the elevation data used for interpolation of a DEM, expressed as an RMS error, we can then use this RMS estimation to generate many elevation models, each with a random element added to it. If we carry out the drainage basin delineation on each DEM, we will get the final result as the sum of all the delineations. Using this method we obtain not just a more reliable drainage basin delineation, but we also get an idea of the confidence of the result. One such example is shown in Figure 9.5, where 100 DEMs have been interpolated from topographic data with 5 m contour intervals. The variability of the data was then estimated to a normal distribution with a standard deviation of ± 5 m. Each of the drainage basin delineations resulted in a raster file coded with 0 for outside and 1 for inside the basin. The sum of all delineation results yields a file with values between 0 and 100, where 0 means outside the basin in all DEMs and 100 means inside the basin in all DEMs.

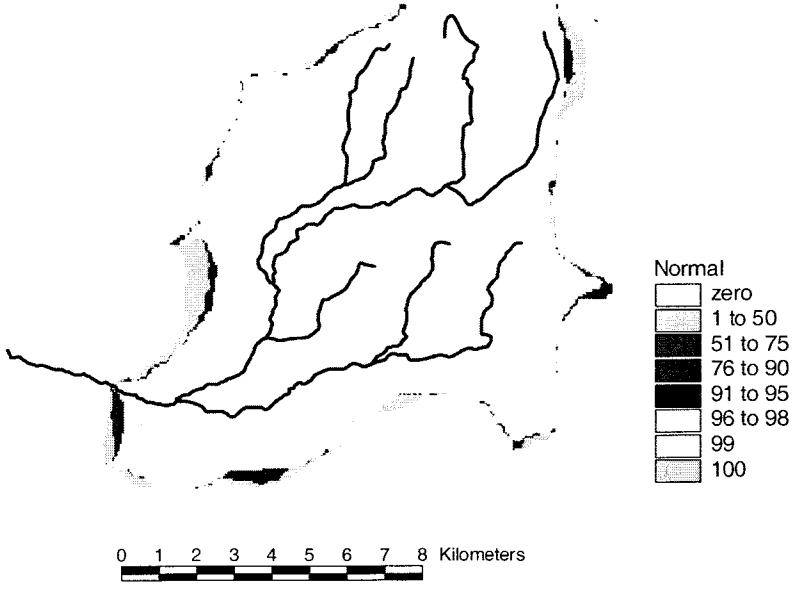


Figure 9.5: Drainage basin delineation based on Monte Carlo simulation. The figure shows the result of 100 basin delineations.

9.2.4 Soil

The important soil characteristics needed for hydrological modelling relate primarily to two aspects of the hydrological processes,

- factors governing the water holding capacity and the movement of water within the soil, and
- factors governing the infiltration capacity and infiltration velocity.

The latter factors have been proven to be of particular importance in the dry tropics, where crusting and sealing of the soil surface can have a large effect on the amount of water actually passing into the soil volume. In hot and dry climates, a slowing down of the infiltration during a rain event will increase the evaporative losses tremendously. The infiltration characteristics of a soil are often very hard to estimate, and fairly rough parameterisation schemes are often used. It is important to point out that infiltration is to a high degree altered by land management and vegetation growth and therefore varies with time. Some examples of infiltration capacities are shown in Table 9.2.

Table 9.2: The effect of different agricultural practices on infiltration capacity (mm/h) (modified from Jones 1997).

Soil type / management	Management / time	Infiltration capacity (mm/h)	Management / time	Infiltration capacity (mm/h)
Silt loam, high organic content:	Under pasture:	27.3	Under cornfield:	6.8
Silt loam, low organic content:	Under pasture:	8.3	Under cornfield:	6.5
Cornfield, unmulched:	In June:	22	In October	6.0
Cornfield, mulched:	In June:	42	In October	39
Coarse soil on sandstone:	Grazed rangeland:	32	Ungrazed:	48
Fine soil on shale:	Grazed rangeland:	18	Ungrazed:	17
	Unimproved pasture:	43	Oak-hickory forest:	76

The factors governing the amount of water that can be held, and the movement of water in the soil are of course difficult to obtain data on, and we have to rely on a limited number of soil properties from which the desired characteristics might be inferred. Soil texture classification, often in the form of fractions of clay, silt and sand, is the single most important data source, and it is also the most commonly available type of soils data. From the texture classification we usually infer estimates of hydraulic conductivity, matric potential and saturated water content. Depth to the underlying rock or other impermeable layer is often desirable, but difficult to obtain.

9.2.5 Climate

A range of different climate variables is needed, and the more advanced model the more demand is put on the climatic data input. For a model to be applied in different environments it is important to restrict the climate data requirements to what is generally available from operational climate stations. But the role of remote sensing for measuring or inferring climate variables is increasing.

The time interval of data input is also a matter for concern. Many detailed models require daily data, but it would often be desirable to use bi-daily data in order to capture the important diurnal variability. A list of common climate variables for a comprehensive modelling of evaporation using the Penman-Monteith formula (see further below) could look like as shown in Table 9.3.

Table 9.3: Example of climate variables needed for evaporation calculations using a Penman-Monteith type of model.

Climate variable	Unit	Source of data
Average daytime temperature	K	observations, max, min and length of day
Cloud cover	%	nearest station or satellite data
Dew-point temperature	K	observations, or min temp. approximation
Average daily wind speed	m/s	nearest station
Precipitation	Mm	observations (snow is inferred from temp.)
Precipitation intensity	mm/h	observations, but often subjective approximations
Relative humidity	%	Observations

Remote sensing based methods for estimating the surface energy balance will play an important role in the near future for our possibility to model climate parameters between observation points. The energy balance plays a very important role for the estimation of evaporation, and substantial progress have been made during the last 10 years to model actual evaporation rates by means of thermal remote sensing. For a detailed discussion of methods, see Kustas *et al.* (1989).

Net radiation can be expressed as the sum of four major components, i.e. the downward and the upward short- and longwave radiation components,

$$R_{net} = Rs_{\downarrow} - Rs_{\uparrow} + Rl_{\downarrow} - Rl_{\uparrow}$$

The two downward fluxes are usually estimated from point observations, which can be extended over relatively large areas (100 km^2 for shortwave and 10 km^2 for longwave under stable weather conditions (Kustas *et al.* 1989), while the two upward components can be inferred by varying success from remote sensing data.

The shortwave reflected radiation, the albedo, can be estimated by means of remote sensing methods (Otterman and Fraser 1976; Wanner *et al.* 1995). Two problems, though, are associated with remote sensing methods. Firstly, remote sensing systems (e.g. Landsat TM or NOAA AVHRR) only measures about 50 per cent of the solar spectrum, and secondly albedo is often quite variable in different angles, whereas most remote sensing methods only measure in one angle. The emitted longwave radiation can be estimated remotely by means of sensing systems in the thermal IR region, typically 10–12 μm. The approach was successfully demonstrated by Jackson (1985), but more research is needed within this field.

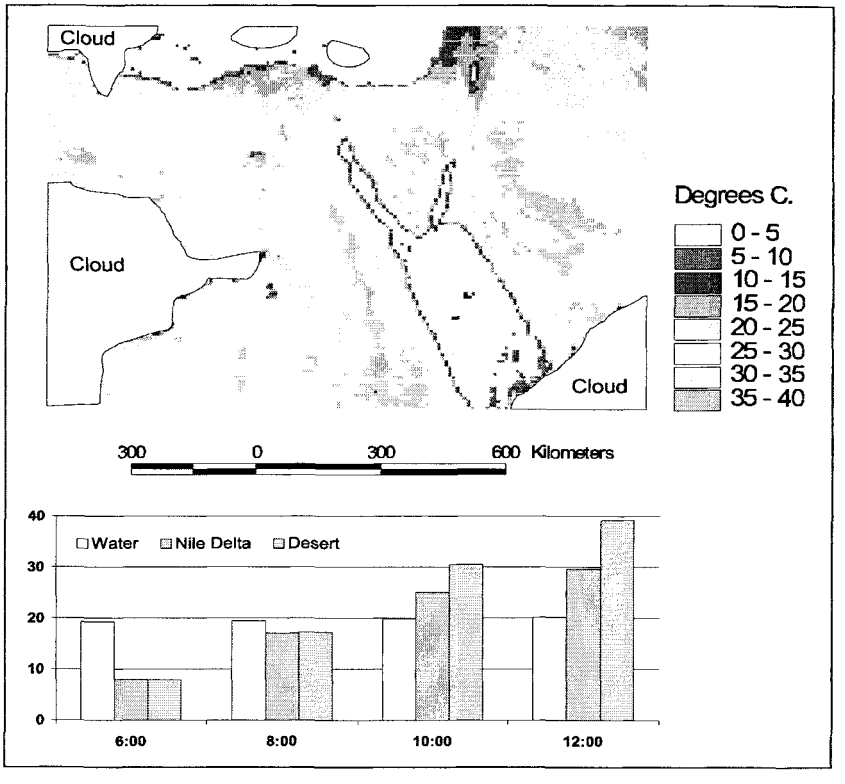


Figure 9.6: The Nile valley in Egypt seen from Meteosat on 1 May 1992. The image shows temperature change (degrees C) from 6 am to 12 noon.

An important concept in modelling evaporation, particularly in dry and hot climates, is thermal inertia, which can be measured by means of temperature change over time. Satellite systems available for thermal inertia studies are e.g. NOAA AVHRR sensor for diurnal changes and the geostationary weather satellites

(like Meteosat) for almost continuous temperature measurements. The diurnal temperature change of the land surface is to a large extent governed by the evaporation of water, particularly in arid and semi-arid environments (insolation effects not considered), and by continuously measuring the diurnal temperature variation (a function of the thermal inertia) there is a good hope of being able to estimate the actual rate of evaporation. More research as well as improved satellite sensors are needed. One example of the temperature change during the morning hours, showing the effect of moisture on the thermal satellite data is shown in Figure 9.6.

In Figure 9.6 it is obvious that the moisture availability along the Nile valley has a cooling effect on the temperature compared with the surrounding desert areas.

9.3 THE LAND SURFACE – ATMOSPHERE INTERFACE

9.3.1 Estimation of potential evaporation

Potential evaporation, i.e. the evaporation from an open water surface, is a fundamental process in the hydrological cycle. A number of different approaches to partition precipitation into evaporation and input to the terrestrial ecosystem have been employed over the last centuries. In the early nineteenth century, John Dalton formulated the law that bears his name and states that evaporation, E, can be expressed as:

$$E = u \cdot (e_s - e_a) \qquad (9.9)$$

where
u = function of wind speed
e_a = the current vapour pressure
e_s = the saturation vapour pressure at that temperature.

Dalton assumed that over water or a moist surface, the air immediately above it rapidly becomes saturated and prevents further evaporation. That is why the wind speed is important. This approach really only describes the potential evaporation, i.e. free evaporation from a constantly wet surface. The actual evaporation on land depends on a range of factors: the depth to the water table, the porosity of the soil and the local heat budget. Transpiration is controlled by a number of additional factors related to the vegetation. From a hydrological point of view, we often add the evaporation and the transpiration to one variable, called evapotranspiration, ET.

Since the formulation of Dalton's law in 1802, a great number of formulae for the calculation of evaporation have been developed. What determines the type of evaporation formula to apply depends to a large extent on the data availability.

The simplest methods are the temperature based ones, where the data needed usually are restricted to temperature, and in some cases sunshine hours. The most widely used formula of this kind is by Thornthwaite (Thornthwaite and Mather 1955), expressing monthly evaporation (cm) as:

$$E = 1.6 \cdot d \cdot \left(\frac{10 \cdot \bar{t}}{I} \right)^{a} \qquad (9.10)$$

where
d = total monthly daylight hours/360
t = mean monthly temperature
$I = \Sigma\, i$ where $i = t^{1.514}$
a = empirical coefficient.

Another approach, more physically based and which corresponds closely to Dalton's law, is the mass transfer approach, where the data requirements also include wind speed and some measure of air humidity. The energy budget approach is based on the assumption that all surplus energy from the net heat balance is used for evaporation. The most sophisticated formulae for calculation of potential evapotranspiration are the combined formulae, which are combinations of the mass transfer and the energy balance approaches. The Penman formula from 1948 (Penman 1948) has been one of the most widely used methods.

The Penman formula considers all the important components of the evaporation process from a soil or water surface, but keeps the data requirement to readily accessible meteorological variables. The formula has been widely used all over the world, but it has also been modified to make the best use of available data. The basic form of the Penman formula is:

$$E = \frac{\Delta \cdot R}{\Delta + \gamma \cdot \rho \cdot L} + \frac{\gamma}{\Delta + \gamma} \cdot f(W) \cdot (1 - RH) \cdot e_a \qquad (9.11)$$

where
Δ = rate of change of saturated water vapour pressure with temperature
γ = psychromatic constant (fairly constant around 0.6)
ρ = density of water
R_n = net radiation
RH = relative humidity
e_a = actual water vapour pressure.

The wind function f(W) is calculated according to the formula below:

$$f(W) = a \cdot (1 + b \cdot W) \qquad (9.12)$$

where
a = constant often set to 0.13 (mm/day, mb)
b = constant often set to 1.1 (s/m) over land and 0.5 over water (Kuzmin 1961)
W = wind speed (m/s).

The first term in the Penman equation is the energy balance, in which the energy available for evaporation ignores heat storage. The net radiation, R is usually calculated from the energy balance equation according to:

$$R = (1-\alpha)\cdot R_s + (1-\alpha_l)\cdot R_{l\_in} - R_{l\_out} \qquad (9.13)$$

where
α = albedo (subscript l means longwave)
R_s = incoming shortwave radiation
R_{l_in} = incoming longwave radiation
R_{l_out} = outgoing longwave radiation .

A further refinement of the Penman equation is the Penman-Monteith formula (Monteith and Unsworth 1990), in which the role of vegetation has been further elaborated. The extra variables taken into consideration are aerodynamic resistance, and is considered in the f(W) term, and the stomata resistance which controls the diffusion of water through the leaf. These variables are reasonably well known for agricultural crops, but for natural ecosystems it is difficult to find appropriate values.

The Penman-Monteith equation has been used and modified by many scholars. One form of the equation is presented here (Monteith 1981):

$$\lambda E = \frac{1}{\Delta + \frac{\gamma\cdot(r_{sfc}+r_{aero})}{r_{aero}}}\cdot\left[(R_n - G)\cdot\Delta + \frac{\rho\cdot c_p\cdot(e_s - e)}{r_{aero}}\right] \qquad (9.14)$$

where
λE = the average daytime latent heat flux (evapotranspiration) (W/m^2)
λ = latent heat of vaporization (J/kg)
E = moisture mass flux (kg/s)
γ = psychromatic constant (Pa/K)
(e_s-e) = the average daytime vapour pressure deficit between a saturated surface at the temperature of the air and the ambient vapour pressure (Pa)
ρ = the atmospheric density (kg/m^3)
c_p = specific heat of water (J/kg, K)
r_{sfc} = surface resistance to evaporation (s/m)
r_{aero} = atmospheric resistance to evaporation, (s/m)
$(R_n - G)$ = net radiation minus ground heat flux (W/m^2).

The role of vegetation is introduced in the calculation of the atmospheric and surface resistances to evaporation, which are functions of the vegetation's leaf area index, root distribution and height. To calculate the atmospheric resistance we need to have information on wind speed and vegetation canopy height, while the surface resistance also needs information on leaf area index. A number of different formulae have been suggested to relate vegetation characteristics (height and LAI) to the surface and aerodynamic resistances, and the reader is referred to Monteith

and Unsworth (1990) and Guyot (1998) for a comprehensive discussion on this issue. Some guidelines for the values of the canopy resistances are the following: for wet vegetation, $r_{sfc} = 0$ (i.e. the Penman-Monteith Equation is equal to the Penman Equation), for a ley crop, r_{sfc} can be between 10 and 30 (s/m) (Bengtsson 1997) and for a pine forest Lindroth (1984) found a value of $r_{sfc} = 100$ as appropriate.

9.3.2 Estimation of actual evaporation and evapotranspiration

In order to estimate potential evapotranspiration as well as actual evapotranspiration, we need either direct flux measurements or it may be possible to infer this from remote sensing measurements. It is also possible to modify the potential evapotranspiration models above by a factor describing the moisture availability. A number of approaches have been taken here.

The first the most simple one is the so called bucket model (Budyko 1969), in which the terrestrial ecosystem was treated as a bucket with a maximum capacity equal to field capacity, FC. The bucket fills when precipitation exceeds evapotranspiration and surface runoff happens when the water level exceeds the field capacity. The actual evaporation is then estimated as a function of the amount of water in the bucket. When the soil water content reach wilting point, WP, the evapotranspiration also becomes zero.

$$Et = m \cdot \lambda E \qquad (9.15)$$

where
Et = actual evapotranspiration
m = moisture availability.

Scaling of the m parameter is mainly a function of the water holding properties of the soil and one simple scaling scheme (Bergström 1992) is shown in Figure 9.7.

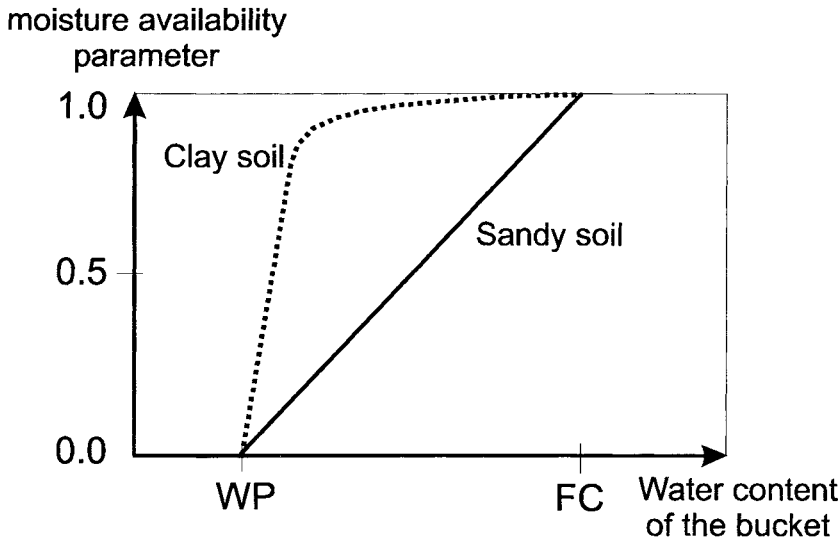


Figure 9.7: The scaling of the moisture availability parameter as a function of soil water content for two different soil types.

The bucket model can be made quite realistic if the evapotranspiration function considers both bare soil evaporation and transpiration through a vegetation canopy, but the real weakness is that the vegetation is still treated as a static component.

The second generation of these models is often referred to as Soil-Vegetation-Atmosphere-Transfer models (SVAT), and the main difference from the first generation models is that vegetation is modelled as dynamic entity. The important vegetation parameters needed for this kind of models is stomata/canopy conductance. Conductance is usually inferred from the absorption of PAR, i.e. photosynthesis rate, which makes sense since water escapes the leaves during opening of stomata. This type of model is sometimes referred to Green SVAT models and for a comprehensive treatment see Sellers (1992). The weakness of these models is that there is no spatial interaction between neighbouring areas, and this is where a spatially distributed hydrological model comes in. The development of green SVAT models coupled to a spatially distributed model for the movement of water on and under the surface, which is still very much at the research stage, is sometimes referred to the 'wetting' of SVAT models.

SVAT models of different degrees of complexity are used in climate models, where there is a coupling between the climatic components of the model and the terrestrial ecosystem through the SVAT model. SVAT models are not yet in operational use in hydrological models, but we anticipate a rapid development within this field.

9.4 THE DISTRIBUTION OF SURFACE FLOW IN DIGITAL ELEVATION MODELS

The amount of water transported as Hortonian overland flow depends, apart from the difference between the precipitation intensity and infiltration capacity, on the potential storage on the ground (Bengtsson 1997). The storage capacity is equal to the volume of small depressions, often called sinks or pits, creating isolated small drainage basins over the surface. The volumes of these sinks can often be estimated by the use of a standard GIS.

The velocity of the overland flow depends on the slope of the surface, the surface roughness (symbolized by a friction factor, f), and the water depth (Bengtsson 1997). The relationship between flow velocity, surface roughness, flow depth and slope can be expressed by Manning's equation:

$$v = M * R^{2/3} * I^{1/2} \qquad (9.16)$$

where v is the water flow velocity, M is Manning's number describing the surface roughness, R is water flow depth, and I is slope. M varies with different water depth, especially for limited depths on vegetated surfaces (Bengtsson 1997).

Surface roughness can be estimated by the use of remote sensing (land cover classification) and the water depth is estimated by combining precipitation intensity, evapotranspiration, infiltration capacity etc. Slope values will normally be estimated from digital elevation models.

Apart from slope, other topographic parameters like aspect and drainage area are also important in hydrological modelling. The aspect value can, in combination with slope, be used for estimation of irradiance over the surface, and drainage area are commonly used as an independent variable in estimation of soil wetness (Beven and Kirkby 1979). Also the distribution of water over a surface is of course critical in distributed modelling. The following section, based on Pilesjö *et al.* (1998), presents one approach to estimate these parameters.

Catchment topography is critical for models of distributed hydrological processes. Slope controls flow pathways for surface flow, and influences the subsurface flow pattern substantially. The key parameter in catchment topography is flow distribution, which tell us how overland flow is distributed over the catchment area.

Stating flow distribution over a land surface is a crucial measurement in hydrological modelling, the use of Digital Elevation Models (DEM) has made it possible to estimate flow on each location over a surface. Based on the flow distribution estimation on each location represented by a DEM, the drainage pattern over an area, as well as various hydrological parameters, such as catchment area and up-stream flow accumulation, can be modelled.

One common approach for measuring flow distribution is the hydrological flow modelling methods that are widely applied to geomorphological and hydrological problems (Moore *et al.* 1994). This method is based on the following basic principles:

1) A drainage channel starts from the close neighbourhoods of peaks or saddle points.
2) At each point of a channel, hydrological flow follows one or more directions of downhill slopes.
3) Drainage channels do not cross each other.
4) Hydrological flow continues until it reaches a depression or an outlet of the system.

One critical and most controversial assumption of the hydrological flow modelling method is the determination of flow direction (or drainage path). In the early development, it was assumed that flow follows only the steepest downhill slope. Using a raster DEM, implementation of this method resulted in that hydrological flow at a point only follows one of the eight possible directions corresponding to the eight neighbouring grid cells (Mark 1984; O'Callaghan and Mark 1984; Band 1986; ESRI 1991). Here we call this approach the 'eight-move' algorithm. However, for the quantitative measurement of the flow distribution, this over-simplified assumption must be considered as illogical and would obviously create significant artefacts in the results, as stated by Freeman (1991), Holmgren (1994), Wolock and McCabe (1995), and Pilesjö and Zhou (1996). More complex terrain is supposed to yield more complicated drainage patterns.

Attempts to solve the problem have led to several proposed 'multiple flow direction' algorithms (Freeman 1991; Quinn *et al.* 1991; Holmgren 1994; Pilesjö 1992; Pilesjö and Zhou 1996). These algorithms estimate the flow distribution values proportionally to the slope gradient, or risen slope gradient, in each direction. Holmgren (1994) summarizes some of the algorithms as:

$$f_i = \frac{(\tan \beta_i)^x}{\sum_{j=1}^{8} (\tan \beta_j)^x}, \text{ for all } \beta > 0 \quad (9.17)$$

where i,j = flow directions (1...8), f_i = flow proportion (0...1) in direction i, $\tan \beta_i$ = slope gradient between the centre cell and the cell in direction i, and x = variable exponent.

By changing the exponent (x) in Equation (9.17), two extreme approaches in estimating flow distribution can be observed. While $x = 1$, flow will be distributed to downhill neighbouring cells proportionally to the slope gradients, as suggested by Quinn *et al.* (1991). The other extreme is when $x \rightarrow \infty$, which will approach towards the 'eight-move' drainage distribution mentioned above. Holmgren (1994) suggested an x value between 4 and 6. This gives a result between a very homogeneous flow distribution when $x = 1$, and a distinctive flow which occurs when x becomes greater than 10.

This approach focuses on the estimation of flow distribution over a raster DEM. A raster DEM is by far the most commonly used elevation models for modelling purposes, and is easily integrated with other types of spatial data. Given the limitation and problems of the 'eight-move' algorithm, a 'multiple flow direction' approach based on analysis of individual surface facets was taken in this

study to estimate flow distribution over the raster DEM. The results from this form-based algorithm were then evaluated and compared with those from the 'eight-move' algorithm.

A number of assumptions when modelling with raster DEM are required in order to isolate the mechanism of overland flow distribution from soil, vegetation and atmospheric impact. The assumptions were:

- The flow from a grid cell to its neighbouring cells is dependent upon the topographic form of the surface.
- Water is evenly distributed over the grid cells (i.e. homogeneous precipitation).
- Every cell in the raster DEM, except sinks, contributes a unit flow and all accumulated flow distributes to its neighbourhood.
- The infiltration capacity over the surface is set to zero.
- The surface is bare (e.g. no vegetation).
- The evaporation is set to zero.
- The topographic form of a cell can be estimated by the use of its eight neighbours (a 3x3 window).

The procedure of a proposed method for the estimation of flow distribution, and subsequently the flow accumulation, from a raster DEM is outlined below:

- Examine each grid cell and the surface facet formed by the 3x3 cell window to classify the facet into 'complicated', 'flat' and 'undisturbed' terrain.
- These three categories are treated separately in determining their flow distribution using the algorithms reported below.
- Results from different categories of the facets are then merged to create flow distribution, and flow accumulation, over the entire study area.
- Classify each facet into 'complicated', 'flat' or 'undisturbed' category.
- The classification of the facet into one of the three categories is based on a normal 3x3 cell filtering process, as outlined below.
- If all the cells in the 3x3 window have the same elevation values the centre cell is marked as a flat surface.
- The eight neighbour cells around the centre cell are examined (clockwise) in order to determine if there are multiple valleys. A valley in this context is defined as one, or more, lower (in elevation) neighbour cells surrounded by cells with higher elevation. The centre cells in these facets are marked as 'complicated' terrain. An example of a facet classified as complicated is presented in Figure 9.8.
- Cells classified as neither 'flat' nor 'complicated' are marked as 'undisturbed'.

90	80	110
110	100	80
90	70	60

Figure 9.8: An example of a surface facet to be classified as 'complicated' terrain. The numbers in the cells denote the elevation values of the centre of the cells. The upper and the upper left cell represent one valley, and the four cells below and right of the centre cell represent another valley (from Pilesjö *et al.* (1998)).

9.4.1 Estimation of topographic form for 'undisturbed' surface facets

There is no absolute way to determine convexity and concavity of the centre cell in a 3 by 3 cell facet. The possible complexity of the surface often implies approximations. One way to approximate is to use a trend surface based on the elevation values of all cells in the facet. Below a method based on a least-squares approximated second-order trend surface (TS) was proposed:

$$TS(x_i, y_i) = a_0 + a_1 x_i + a_2 y_i + a_3 x_i^2 + a_4 y_i^2 + a_5 x_i y_i \quad (9.18)$$

where

$i = 1, \ldots, 9$ = The index numbers of the centre cell and its eight neighbours.

$a_0, \ldots, a_5$ = The constants for the second-order trend surface.

x_i, y_i = The cell co-ordinates (central cell) in a local system.

Different forms of surfaces should not be modelled in the same way in terms of flow distribution. Complicated terrain, characterized by a number of valleys and ridges, has to be investigated mainly in order to obtain different possible water channels from the centre cell to its neighbours. The flow distribution over flat areas is obviously dependent upon the surrounding topography.

The presence of 'complicated' terrain and flat areas is relatively rare. In most surface facets on a DEM the topography can be classified as 'undisturbed'. This means that the surface is characterized by one single topographic form: concavity or convexity. Convex forms normally occur in the higher parts of the terrain, while the valleys are characterized by concave topographic forms. Since the flow distribution over a convex surface is divergent, and the flow converges over a concave surface, it seems more appropriate to include the topographic form when estimating flow distribution.

The proposed method has demonstrated its ability to produce a better simulation for flow accumulation compared with the 'eight-move' algorithm which is commonly used in today's GIS. However, the results have to be statistically tested, by applying the new algorithm to mathematically generated surfaces.

Another challenge in future research will be focused on appropriate algorithms for generating drainage network, appropriate treatment of sinks, and introduction of more environmental variables into the hydrological modelling processes.

9.5 ESTIMATION OF SUBSURFACE FLOW

Subsurface flow can be described by the use of Darcy's law:

$$q = K * I \qquad (9.19)$$

where q is water flow velocity, K is hydraulic conductivity, and I is the pressure gradient. Since water only can flow in pores filled with water, the flow velocity in these pores, v_p, is

$$v_p = q / \theta \qquad (9.20)$$

where θ is soil moisture. At a point at a specific level, z, the water pressure is ψ.

The total pressure (H), or height of water table, is then

$$H = z + \psi \qquad (9.21)$$

Subsequently, the pressure gradient I = -dH/dz incorporated in Darcy's equation gives:

$$q = -K(1 + d\psi / dz) \qquad (9.22)$$

Subsurface flow can be divided into unsaturated flow in the unsaturated zone and groundwater flow in the saturated zone (see Figure 9.2). When modelling hydrology, the unsaturated flow is often generalized to the vertical dimension, while the saturated flow is modelled in three dimensions. The following two sections are based on Bengtsson (1997).

9.5.1 Estimation of subsurface unsaturated flow

Estimation of unsaturated subsurface flow is normally based on Richard's Equation. The difference between incoming and outgoing flow in a soil with a vertical extension dz corresponds to a difference in soil moisture, θ, in the soil volume during the time dt as:

$$\frac{d\theta}{dt}+\frac{dq}{dz}+e=0 \qquad (9.23)$$

where e = water loss to vegetation in the root zone and where z has the same positive direction as the flow q. The flow can be calculated by the use of Darcy's law (see above). If we exclude e from Equation 9.23, and substitute q according to Darcy's law we get:

$$\frac{d\theta}{dt}=\frac{d}{dz}\left(K+K\frac{d\psi}{dz}\right) \qquad (9.24)$$

Since hydraulic conductivity as well as water pressure depends on soil moisture it is preferable to rewrite equation 9.24 in a way that makes θ the only unknown variable. This can be done with a new variable, D, defined as $D = K*d\psi/d\theta$. We then get:

$$\frac{d\theta}{dt}=\frac{d}{dz}\left(D\frac{d\theta}{dz}+K\right) \qquad (9.25)$$

If we know how D and K vary with θ we can then solve the equation numerically.

However, if the soil profile is homogeneous the soil moisture changes rapidly and discontinuously with depth. It is then easier to handle an equation including water pressure instead of soil moisture, since water pressure changes continuously in the profile. This is possible if we substitute with 'specific soil water capacity', defined as $C(\psi) = d\theta/d\psi$. We then get Richardson's Equation as:

$$C\frac{d\psi}{dt}=\frac{d}{dz}\left(K\frac{d\psi}{dz}+K\right) \qquad (9.26)$$

To solve Equation 9.26 we normally have to work with numerical methods. Besides C and K we also need knowledge about initial values along the soil profile and the water pressure above and below the profile. At the soil surface the pressure can be calculated as a function of the infiltration rate, f, or evaporation rate, $-f$, according to:

$$\frac{d\psi}{dz} = \frac{f}{K} - 1 \quad (9.27)$$

If the surface is saturated the water pressure is equal to the water depth. Below the soil profile, at the water table, water pressure is zero.

9.5.2 Estimation of subsurface saturated flow

The groundwater flow is three-dimensional, and the permeable material is often heterogeneous. The hydraulic properties of the ground are characterized by anisotropy, and the hydraulic conductivity is different in different directions. The flow can be estimated by using Darcy's law:

$$v = K * I \quad (9.28)$$

where v is flow velocity and K is saturated hydraulic conductivity and the pressure gradient I is defined as:

$$I = -\frac{dH}{dx} \quad (9.29)$$

where H is total hydraulic pressure height and x is co-ordinates in the flow direction. As mentioned above, Darcy's law is only valid for relatively slow flow, and is not applicable in cracks and in material coarser than gravel.

Since the water cannot flow through the soil mineral particles, but only in the pores between them, the velocity of the water particles, v_p, is faster than flow velocity according to:

$$v_p = v / n \quad (9.30)$$

where n is the soil porosity.

In distributed hydrological modelling, the water table is normally estimated by the use of a DEM. If the distance from the surface to the groundwater is known for every cell, an initial water table can be estimated. When this is done the saturated flow and the fluctuations in the groundwater table can be modelled according to Darcy's law. By comparison of total water pressure in each cell with its neighbours, and with knowledge about hydraulic conductivity in different directions, hydrological conditions can be modelled in time and space.

9.6 SUMMARY

We have given an overview of problems and solutions related to the development of spatially distributed hydrological models. There is a severe gap concerning

spatial and temporal scales between the processes we need to mimic and the availability of data. Remote sensing is playing an increasing role in filling this gap, but a number of critical research themes exist where the knowledge is not yet adequate. One such important field is the parameterization of the vegetation's role in the evapotranspiration. Interesting approaches based on the use of thermal remote sensing are under way, but there is need for much more research and improved sensing systems. Another such important field is on modelling the transfer of water between the saturated and the unsaturated zone. One thing we are confident about is that modelling in a GIS environment, with a strong link to remote sensing, is a promising way to go.

9.7 REFERENCES

Abbot, M.B. and Refsgaard, J.C. (eds), 1996, *Distributed Hydrological Modelling*. Dordrecht, Kluwer Academic Publishers.

Andersson, M.G. and Burt, T.P., 1985, *Hydrological Forecasting*. Chichester, John Wiley and Sons.

Andersson, U. and Nilsson, D., 1998, *Distributed Hydrological Modelling in a GIS Perspective – An evaluation of the MIKE SHE Model*. Dept. of Physical Geography, Lund University.

Asrar, G. (ed.), 1989, *Theory and application of optical remote sensing*, New York, Wiley.

Band, L.E., 1986, Topographic partition of watersheds with digital elevation models. *Water Resources Research*, **22**, 15–24.

Begue, A., 1993, Leaf area index, intercepted photosynthetically active radiation and spectral vegetation indices: a sensitivity analysis for regular-clumped canopies. *Remote Sensing of Environment*, **46**, 45–59.

Bengtsson, L., 1997, *Hydrologi – teori och processes*. In Swedish. Svenska Hydrologiska Rådet. Institutionen för Tekn Vattenresurslära. Lund University.

Bergström, S., 1992, *The HBV-model, its structure and applications*. SMHI report (Hydrology) 4.

Beven, K.J. and Kirkby, M.J., 1979, A physically-based, variable contributing area model of basin hydrology. *Hydrological Science Bulletin*, **24**, 1–10.

Beven, K.J. and Moore, I.D., (eds), 1993, *Terrain analysis and distributed modelling in hydrology*. Chichester , John Wiley and Sons.

Brady, N.C. and Weil, R.R., 1996, *The nature and properties of soils*. New Jersey , Prentice-Hall Inc.

Budyko, M.I., 1969, The effect of solar radiation variations on the climate of the Earth. *Tellus*, **21**, 611–1044.

Burrough, P.A. and McDonnell, R.A., 1998, *Principles of Geographical Information Systems*. Oxford, Oxford University Press.

Calder, I.R., 1986, A stochastic model of rainfall interception. *Journal of Hydrology*, **89**, 65–71.

Chorley, R.J., 1977, *Introduction to Physical Hydrology*. London, Methuen and Co. Ltd..

ESRI, 1991, *Cell-based Modelling with GRID*, Environmental System Research Institute, Redlands, CA.

Freeman, T.G., 1991, Calculating catchment area with divergent flow based on a regular grid. *Computers & Geosciences*, **17**, 413–422.

Gash, J.H.C., 1979, An analytical model of rainfall interception by forests. *Quaterly Journal of Roy. Met. Soc.*, **105**, 43–55.

Grip, H. and Rhode, A., 1994, *Vattnets Väg från Regn till Bäck*. In Swedish. Forskningsrådens Förlagstjänst, Uppsala.

Guyot, G., 1998, *Physics of the environment and climate*. Chichester, John Wiley and Sons.

Holmgren, P., 1994, Multiple flow direction algorithms for runoff modelling in grid based elevation models: An empirical evaluation. *Hydrological processes*, **8**, 327–334.

Huete, A.R., 1989, *Soil influences in remotely sensed vegetation-canopy spectra*. Chapter 4 in Asrar 1989.

Jackson, R.D., 1985, Evaluating evapotranspiration at local and regional scales. *Proceedings IEEE*, **73**, 1086–1095.

Jones, J.A.A., 1997, *Global hydrology – Processes, resources and environmental management*. Edinburgh, Longman, 390 pp.

Kirby, M.J., Naden, P.S., Burt, T.P. and Butcher, D.P., 1993, *Computer simulation in Physical Geography*. Chichester, John Wiley and Sons.

Kustas, W.P., Jackson, R.D. and Asrar G., 1989, *Estimating surface energy balance components from remotely sensed data*. In: Asrar, 1989.

Kuzmin, P.O., 1961, Hydrophysical investigations of land waters. International Science Hydrology, *International Union of Geodesy and Geophysics*, **3**, 468–478.

Lind, M. and Fensholt, R., 1999, The spatio-temporal relationship between rainfall and vegetation development in Burkina Faso. *Danish Journal of Geography*, 1999, **2**, 43–56.

Lindroth, A., 1984, *Seasonal variation in pine forest evaporation and canopy conductance*. Avhandling 758, Uppsala Universitet.

Lindström, G., Gardelin, M., Johansson, B., Persson, M. and Bergström, P., 1996, *HBV-96-En Areellt Fördelad Modell för Vattenkrafthydrologin*. In Swedish. SMHI RH Nr. 12, SMHIs tryckeri, Norrköping.

Mark, D.M., 1984, Automated detection of drainage networks from digital elevation models. *Cartographica*, **21**, 168–178.

Monteith, J.L., 1981, Evaporation and surface temperature. *Quaterly Journal of the Roy. Met. Soc.*, **107**, 1–27.

Monteith, J.L. and Unsworth, M., 1990, *Principles of environmental physics*. London , Arnold, 285 pp.

Moore, I.D., Grayson, R.B. and Ladson, A.R., 1994, *Digital terrain modelling: a review of hydrological, geomorphological and biological applications* in Beven, K.J. and Moore, I.D. (eds), Terrain Analysis and Distributed Modelling in Hydrology, Chichester, John Wiley and Sons, UK, 7–34.

Myneni, R.B. and Williams, S.E., 1994, On the relationship between fAPAR and NDVI. *Remote Sensing of Environment*, **49**, 200–211.

O'Callaghan, J.F. and Mark, D.M., 1984, The extraction of drainage networks from digital elevation data. *Computer Vision, Graphics and Image Processing*, **28**, 323–344.

Oke, T.R., 1995, *Boundary Climate Layers*. Cambridge, University Press.

Otterman, J. and Fraser, R.S., 1976, Earth-atmosphere system and surface reflectivities in arid regions from Landsat MSS data. *Remote Sensing of Environment*, **5**, 247–266.

Penman, H.L., 1948, Natural evaporation from open water, bare soil and grass. *Proceedings of the Roy. Soc. Of London*, A194, 20–145.

Persson, D.A. and Pilesjö, P., 2000, Digital elevation models in precision farming. Sensitivity tests of different sampling schemes and interpolation algorithms for the surface generation. *Proceedings of the 2nd International Conference on Geospatial Information in Agriculture and Forestry*, Lake Buena Vista, Florida, 9 – 12 January 2000. pp. II-214-221.

Pilesjö, P., 1992, *GIS and Remote Sensing for Soil Erosion Studies in Semi-arid Environments*, PhD thesis, Meddelanden fran Lunds Universitets Geografiska Institutioner, Avhandlingar CXIV.

Pilesjö, P. and Zhou, Q., 1996, A multiple flow direction algorithm and its use for hydrological modelling in *Geoinformatics '96* Proceedings, April 26-28, West Palm Beach, Florida, 366–376.

Pilesjö, P. and Zhou, Q., 1997, Theoretical estimation of flow accumulation from a grid-based digital elevation model, in *Proceedings of GIS AM/FM ASIA '97 and Geoinformatics '97 Conference*, 26-29 May, Taipei, 447–456.

Pilesjö, P., Zhou, Q. and Harrie L., 1998, Estimating Flow Distribution over Digital Elevation Models using a Formed-Based Algorithm. *Geographic Information Sciences*, **4**, 44–51.

Pinter, P.J., 1992, Solar angle independence in the relationship between absorbed PAR and remotely sensed data for alfalfa. *Remote Sensing of Environment*, **46**, 19–25.

Priestly, C.H.B. and Taylor, R.J., 1972, On the assessment of surface heat flux and evaporation using large scale parameters. *Monthly Weather Review*, **100**, 36–55.

Prince, S.D. and Goward, S.N., 1995, Global primary production: a remote sensing approach. *Journal of Biogeography*, **22**, 815–835.

Quinn, P., Beven, K., Chevallier, P. and Planchon, O., 1991, The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. *Hydrological Processes*, **5**, 9–79.

Selby, M.J., 1982, *Hillslope Materials and Processes*. New York, Oxford University Press.

Sellers, P.J., 1992, *Biophysical models of land surface processes in Trenberth 1992*, Climate System Modelling, Cambridge University Press, 780 pp.

Sellers, P.J., Los, S.O., Tucker, C.J., Justice, C.O., Dazlich, D.A., Collats, G.J. and Randall, D.A., 1996, A revised land surface parameterisation (SiB2) for atmospheric GCMs. Part II: the generation of global fields of terrestrial biophysical parameters from satellite data. *Journal of Climate*, **9**, 706–737.

Shaw, E., 1993, *Hydrology in Practice*. London, Chapman and Hall.

Thornthwaite, C.W. and Mather, J.R., 1955, The water balance. Centerton, NJ, Laboratory for climatology publications in *climatology*, **8**, 1–86.

Wanner, W., Strahler, A.H., Muller, J.P, Barnsley, M., Lewis, P., Li, X. and Schaaf, C.L.B., 1995, Global mapping of bi-directional reflectance and albedo for the EOS MODISA project: the algorithm and the product. *Proceedings IGARSS '95*, Firenze,

Wolock, D.M. and McCabe, Jr., G.J., 1995, Comparison of single and multiple flow direction algorithms for computing topographic parameters in TOPMODEL. *Water Resources Research*, **31**, 1315–1324.