Thermal Remote Sensing of Surface Soil Water Content with Partial Vegetation Cover for Incorporation into Climate Models

ROBERT R. GILLIES

Earth System Science Center, The Pennsylvania State University, University Park, Pennsylvania

TOBY N. CARLSON

Department of Meteorology, The Pennsylvania State University, University Park, Pennsylvania

(Manuscript received 14 January 1994, in final form 1 June 1994)

ABSTRACT

This study outlines a method for the estimation of regional patterns of surface moisture availability (M_0) and fractional vegetation (Fr) in the presence of spatially variable vegetation cover. The method requires relating variations in satellite-derived (NOAA, Advanced Very High Resolution Radiometer) surface radiant temperature to a vegetation index (computed from satellite visible and near-infrared data) while coupling this association to an inverse modeling scheme. More than merely furnishing surface soil moisture values, the method constitutes a new conceptual and practical approach for combining thermal infrared and vegetation index measurements for incorporating the derived values of M_0 into hydrologic and atmospheric prediction models.

Application of the technique is demonstrated for a region in and around the city of Newcastle upon Tyne situated in the northeast of England. A regional estimate of M_0 is derived and is probably good for fractional vegetation cover up to 80% before errors in the estimated soil water content become unacceptably large. Moreover, a normalization scheme is suggested from which a nomogram, "universal triangle," is constructed and is seen to fit the observed data well. The universal triangle also simplifies the inclusion of remotely derived M_0 in hydrology and meteorological models and is perhaps a practicable step toward integrating derived data from satellite measurements in weather forecasting.

1. Introduction

The combination of remote sensing measurements and inverse modeling techniques has advanced to the stage where land surface processes can be described in terms of parameters (such as soil water content) whose values are generally representative of the heterogeneity apparent at the earth's surface (Price 1990; Nemani et al. 1993). In fact, the importance of specifying the surface energy budget that forms the lower boundary conditions of mesoscale (meso-β-scale or meso-α-scale) models cannot be overemphasized (Pielke 1984; Benjamin and Carlson 1986; Mahfouf et al. 1987; Pinty et al. 1989; Avissar and Pielke 1989). It would appear then that the problem of specifying the correct partition of available energy between sensible and latent heat (assuming the sensible heat into the ground is negligible) at each horizontal grid point of such models might appear to be solved.

Indeed, current methods of analysis (Soares et al. 1992; Carlson et al. 1993) that reduce remotely derived measurements to physical parameters at the surface generally involve a complex series of processing steps. These begin with the acquisition of data for cloud-free conditions, which is then subjected to a number of necessary processing procedures—georeferencing, calibration, and correction. Even for the National Oceanic and Atmospheric Administration's (NOAA) Advanced Very High Resolution Radiometer (AVHRR), currently the most appropriate multispectral (including thermal) sensor for land surface studies, data reduction is not only time consuming but is also subject to systematic problems associated with the sensor itself. The next stage, which often involves coupling the satellite data to some sort of remote sensing algorithm [soil–vegetation–atmosphere transfer (SVAT) scheme] is similarly cumbersome. The implications of all this processing time is that it is virtually impossible to incorporate (at least in any interactive way) satellite-derived data into existing mesoscale and or climate modeling schemes.

A more serious issue though lies in the theoretical aspects of remote sensing the surface energy balance and the surface soil water content where biophysical factors at the surface impinge upon the measurements, as is the case for thermal infrared radiometry. This results in an ambiguous solution for the energy balance and surface soil water content. The problem lies in the

Corresponding author address: Dr. Robert R. Gillies, Earth System Science Center, The Pennsylvania State University, University Park, PA 16802-5094.

fact that these parameters are not defined in the presence of vegetation as the surface radiant temperature depends upon the radiances both from vegetated surfaces (e.g., leaves) and from the soil surface—the so-called partial canopy.

This paper therefore proposes a method for dealing with partial vegetation canopies in a manner that allows the resolution of soil *surface* moisture availability M_0.[1] The ambiguity due to vegetation is treated in two ways: by making use of an additional parameter, the normalized difference vegetation index (NDVI), which serves as a surrogate for the vegetation amount, and by imposing physical constraints on the data. These constraints are based upon an interpretation of the NDVI surface radiant temperature scatterplot, which is then further interpreted using a SVAT scheme. The second development takes this algorithm one step further and substitutes a normalization scheme to obviate many of the intermediary steps involved in reducing the remotely sensed data.

Accordingly, this paper begins by presenting the analysis elements [vegetation index (NDVI) and surface radiant temperature] of an inversion technique (section 3) that subsequently reduces the data to quantities of surface water content (surface moisture availability) and fractional vegetation cover. The analysis uses local area coverage (LAC) data (section 2) obtained from the NOAA satellite platforms for the city of Newcastle upon Tyne (including the surrounding hinterland), situated in the northeast of England. Section 4 then addresses the temporal variability of the AVHRR observations, and so in the SVAT simulations, with a procedure that accounts for the external effects of the atmosphere and terrain in deriving the surface water content. Finally, section 5 concludes with a discussion of the refinement of the technique, its future verification, and reflects upon applications of the method.

2. Study site and data

a. Description of the study area

There is no a priori reason or particular justification for selecting Newcastle upon Tyne as the area of study—any conurbation in Britain would presumably suffice. Principal criterion is simple in the sense that the location be a reasonably discrete urban setting with ample regional variability in land cover. The degree of cloud cover is a particularly limiting factor in a program of this nature; however, weather systems tracking across Britain tend to have a high degree of periodicity. For this reason, no region necessarily exhibits a significantly higher frequency of clear skies than any other. A final consideration is to avoid the incorporation of data that is a view of particularly steep vertical relief since this raises a series of issues that are cumbersome to deal with, if not a serious source of error (Carlson 1986).

b. Data

Multiband radiometer measurements were obtained from the NOAA AVHRR provided by the department of Applied Physics and Electronic and Manufacturing Engineering at the University of Dundee in Scotland. Four AVHRR images were acquired at approximately 1315 UTC 21 May 1989, 1336 UTC 1 June 1985, 1400 UTC 5 July 1989, and 1 August 1990 over Northern England including Newcastle upon Tyne. The imagery was calibrated (at the full 10-bit resolution) using the algorithms contained in the NOAA technical memorandum (Lauritson et al. 1988) and atmospherically corrected for the thermal channels using a radiative transfer model (Price 1983) through application of vertical profiles of pressure, temperature, and moisture from radiosonde data collected at 1200 UTC from the most meteorologically representative station. Thus the radiance data from channels 4 and 5 (10.3–11.3 μm; 11.5–12.5 μm) are at-surface values of radiant temperature, whereas channels 1 and 2 (0.58–0.68 μm; 0.725–1.1 μm) are at-sensor (exoatmospheric) values of reflectance. Each overpass was rectified to the U.K. national grid with precise georeferencing to ensure that from scene to scene each resampled (using a nearest neighbor algorithm) 1-km pixel represented the exact same location. An evaluation of the registration procedure indicated a geometric accuracy where the maximum root-mean-square error (rmse) was 0.8 km.

Landsat TM data was also utilized in this study. These data were acquired from the National Remote Sensing Centre Ltd, Farnborough, England. The digital number data (8 bit) was converted to exoatmospheric values of reflectance according to Markham and Barker (1986) and registered to the U.K. national grid using a cubic convolution algorithm. After resampling, the pixel dimensions were 25 m in both the x and y directions.

3. Elements of the satellite observations and method of reduction

a. The temperature–vegetation index scatterplot

The surface radiant temperature depends on the soil water content and the distribution of vegetation. The former is inferred directly from remote sensing, whereas the latter has to be derived in some appropriate way, that is, as a vegetation index (Pearson and Miller 1972; Gutman 1987; Huete 1988; Pinty and Verstraete 1992). Nemani and Running (1989), Carlson et al. (1990), and Carlson et al. (1994) demonstrate the physical relationship between vegetation index (NDVI) and surface radiant temperature T_0 in the form of a scatterplot, the implication of which is that the sensi-

[1] The surface moisture availability is defined as the ratio of the soil water content to that at field capacity.

tivity of the surface radiant temperature to surface soil water content differs for the leaf and for the soil surface around the plants. This sensitivity tends to be much greater for the bare soil as a function of soil water content (surface or root zone) than for the leaves. This association between temperature and vegetation index therefore permits the resolution of a more representative surface soil water content than that previously obtainable by other inversion techniques (Carlson 1986; Soares et al. 1988; Diak 1990; Soares et al. 1992) that neglect the effect of a spatially variable vegetation cover.

Figure 1 shows just such scatterplots, as derived from the NOAA AVHRR for the Newcastle area for four dates (21 May 1989, 1 June 1985, 5 July 1989, 1 August 1990). Physical constraints are imposed on these scatterplots that establish (a) a locus of points indicative of zero water content (a leading "warm edge" observed in all four cases); (b) zero vegetation cover ($NDVI_0$), that is, bare soil and corresponding to the base of the distribution at low NDVI; and (c) full (100%) vegetation cover ($NDVI_s$). These three features set the domain for the SVAT model simulations as will be described later. In the meantime, however, we elaborate upon the significance and determination of the warm edge, $NDVI_0$, and $NDVI_s$.

These three features are not arbitrary but have physical meaning. If the observations contain a complete spectrum of vegetation amount and soil water content, then by definition the scatterplot must be bounded by zero soil water content and by zero vegetation cover, as well as by 100% vegetation cover (no bare soil visible to the radiometer) and by field capacity. In fact, the warm edge of the distribution is consistently well defined and sharply bounded. It constitutes a limit to the warmest and therefore, by implication, the driest pixels in a mixture of bare soil and vegetation. The distribution is characterized by two limits at high and low NDVI. At high NDVI ($NDVI_s$) there exists a rounded top to the distribution, which is identified with a nearly maximum amount of reflection from the 100% vegetation cover. The relatively wide, and flat base of the distribution ($NDVI_0$) is taken as corresponding to sunlit bare soil. The cool side of the distribution would also be well defined except that it is subject to the greatest contamination by clouds, water bodies, and shading by surface elements and so tends to be ragged.

1) THE WARM EDGE

A warm edge is presumed to exist when the warm side of the distribution in the scatterplot is both sharply defined and smoothly varying. It is important, however, to determine without conjecture the form and position of the edge. To achieve this, an algorithm was designed to extract the data points that lie within a discretionary

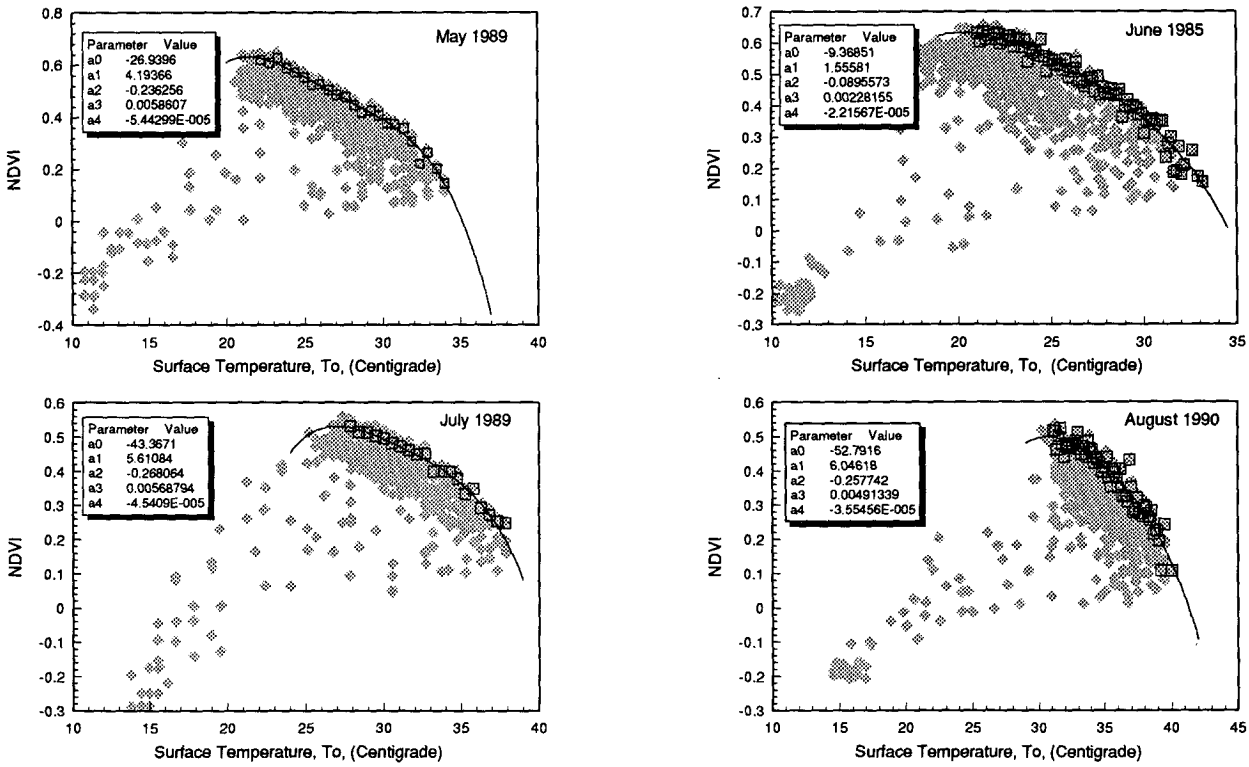

FIG. 1. NDVI versus surface radiant temperature T_0 for all four cases as labeled. In each case the extracted warm-edge pixels are depicted along with the analytical form of the edge; the polynomial coefficients for the edge are also given.

range of the maximum value of NDVI encountered for a narrow temperature band (rounded to the nearest 0.5°C). The set of values of NDVI obtained for each range of temperature are averaged (denoted in Fig. 1 as unfilled squares) in order—this reduces bias in the regression due to variations in the frequency of data points in any given area of the NDVI–T_0 scatterplot.

The next step is to represent the warm edge mathematically by fitting the averaged data to a regression equation. A fourth-order polynomial reproduces the features as shown by Fig. 1. The equation of the line forms a maximum (which is associated with $NDVI_s$), borders the warm edge, and if necessary can be extrapolated to intercept $NDVI_0$.

2) Bare soil NDVI—$NDVI_0$

$NDVI_0$ is not a constant in time or space; its value is determined by small-scale (1 m or less) properties of the surface, such as chemical composition, organic constituents, soil moisture, and roughness. At larger geophysical scales (10 m–1 km) topography will—through changes in slope, aspect, and elevation—alter the illumination of the soil and so alter the value of $NDVI_0$.

At coarser resolutions like that of the AVHRR, the determination of a representative $NDVI_0$ generally requires recourse to higher-resolution visible and near-infrared measurements (e.g., Landsat TM measurements), from which it is possible to establish an estimate[2] value for $NDVI_0$, for which purpose Fig. 2 is relevant. For visible/near-infrared scatterplots of this type (typical semivegetated scenes), the data often take the form of a triangle whose base (right-hand side) consists of a nearly straight line emanating from approximately the origin. This is commonly referred to as the "line of soils" (Jackson 1983). Observations (Baret et al. 1993) and corroborating theoretical soil lines (Jasinski and Eagleson 1990) infer that variation in the range in values of soil line parameters is limited. This implies the existence of a "global" soil line. To this end, the soil line (lower limit of vegetation) is approximated by a simple linear regression

$$a_{nir} = ma_{vis} + c, \tag{1}$$

where a represents the reflectance measurements of the near-infrared and visible channels of the TM sensor, respectively, subscripts nir and vis. The slope (gain) of the line is denoted by m and c is the intercept (offset).

NDVI is formulated as

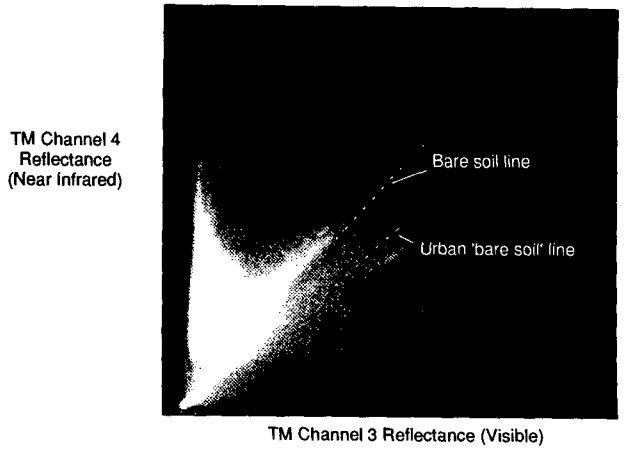

FIG. 2. Landsat TM channel 4 reflectance versus channel 3 reflectance scatterplot. The differences in the spectral signatures of urban surfaces and bare soil (nonurban) are illustrated.

$$NDVI = \frac{a_{nir} - a_{vis}}{a_{nir} + a_{vis}}. \tag{2}$$

Substitution of Eq. (1) in Eq. (2) expresses NDVI along the line of soils in terms of the visible reflectance only

$$NDVI_0 = \frac{(m-1)a_{vis} + c}{(m+1)a_{vis} + c}. \tag{3}$$

If the line of soils can be extrapolated through the origin ("global" soil line), the intercept c is 0 and Eq. (3) becomes effine:

$$NDVI_0 = \frac{m-1}{m+1}, \tag{4}$$

the value of which is independent of a_{vis}, that is, a constant for all points on the line of soils.

A fundamental problem arises in the definition of $NDVI_0$ with this dataset since the study region includes both built-up (urban) and rural areas. $NDVI_0$ has to be resolved since this value sets the lower limit for the domain over which the SVAT simulations are made. To determine between "soil lines," the two land surface types (urban and rural) were separated and plotted independently as in Fig. 2. This determination is because bare soil and built-up surfaces are not similar spectrally (at least in a visible–infrared domain); thus, there remains the potential of discerning two values for $NDVI_0$. Utilizing Eq. (3), it is possible to approximate a value of $NDVI_0$ over the frequency distribution range of a_{vis} within which most of the data points lie. As plotted in Fig. 3, and overlaid with the frequency distribution of the surface's pixels, it follows, respectively, that approximate values for $NDVI_0$ for the urban surface and for actual bare soil are 0.08 and 0.23—the lower of the two defines $NDVI_0$ for this analysis.

[2] It is important to note that the band centers of the two instruments (Landsat TM; *NOAA-11*, AVHRR) are not identical. Exact correspondence of the results for $NDVI_0$ is therefore not to be expected, even from simultaneous measurements. Nevertheless, it is possible to derive similar values for NDVI from the different sensors (Price 1987).

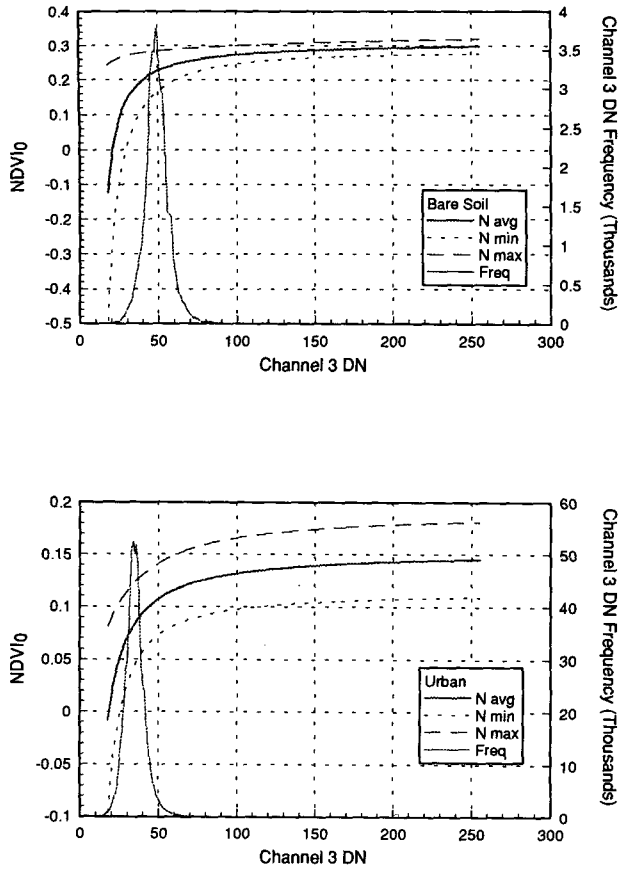

FIG. 3. Solution to Eq. (3) for urban and nonurban surfaces showing NDVI versus channel 3 DN. Curves labeled maximum and minimum refer to the uncertainty in measuring the slope and intercept in Fig. 2. The distribution of DN values is overlaid. $NDVI_0$ is taken as an average over the range of NDVI defined by half the height of the DN frequency distribution. The DN to reflectance conversion is achieved using the method of Markham and Barker (1986).

3) SATURATION NDVI—$NDVI_s$

Representative at-sensor values for $NDVI_s$ for most AVHRR datasets (Tucker 1986; Price 1990) lie at around 0.6. Furthermore, radiative transfer calculations suggest that NDVI increases only marginally thereafter with increasing vegetation amount (increasing leaf area index) at 100% cover. Values lying much below this maximum (even considering errors in calibrations and/or the condition of the vegetation) are unlikely to represent 100% cover (Asrar et al. 1984). To this end, $NDVI_s$ is defined as that point at or above 0.6 where the warm edge *begins* to bend horizontally. Given this criteria, the reader will notice that $NDVI_s$ (Fig. 1) is not present on two days—July 1989 and August 1990. The reason for this is not inherently clear but given that the calibration and corrections to the data are sound, presumably the measurements of NDVI for the dates in question are indicative of a fractional vegetation cover less than 100%. As a precursor and necessary element to the SVAT simulations, the vegetation cover fraction corresponding to the observed maximum in NDVI has to be resolved. There are two ways to approach this problem. As will be shown later, an integral part of the SVAT simulations is to resolve an association between fractional vegetation cover (Fr) and NDVI. This relationship, which is highly nonlinear, could be used to assess Fr for both days. The basis for determining this relationship (i.e., May 1989 and June 1985), however, is inappropriate because these are a limited set of observations and so, it would not be possible to state with reasonable reliability that it would necessarily hold for other cases. In the meantime and certainly until more cases are analyzed, the estimation of Fr for July and August was based upon a linear interpolation between NDVI and Fr between $NDVI_0$ and $NDVI_s$, which was taken to be 0.62. This is quite acceptable because we are only interested in establishing a fractional vegetation cover estimate for the maximum *observed* NDVI for the two days, which was 0.52 (Fig. 1). In making this assumption the associated errors in determining Fr will be small.

b. The SVAT model—Overview and hypotheses

The SVAT model, a key link in the reduction of the observed and corrected satellite measurements, is generic in the sense that models of similar type (i.e., one-dimensional boundary layer models with a vegetation parameterization scheme) are equally pertinent to such inversion procedures (the exact form being unimportant). The SVAT model used here constitutes work and development since its original construct (Carlson and Boland 1978), with subsequent improvements in the parameterization of the boundary layer (Carlson et al. 1981), to the inclusion of the vegetation subscheme (Taconet et al. 1986; Lynn and Carlson 1990).

A quantitative interpretation of the scatterplots (Fig. 1) is made using the SVAT model within the constraints imposed by the warm edge, $NDVI_0$, and $NDVI_s$. This is accomplished by varying two SVAT model parameters, that is, those most directly responsible for the observed variation in surface radiant temperature and vegetation index—the fractional vegetation cover and the surface moisture availability M_0. This is done for each image by setting initial conditions and running the model over a 24-h period. The procedure presupposes as before (Carlson et al. 1994) the following: (a) that the warm edge is the limiting value for the soil water content, which is equivalent to a dry soil surface; (b) there is a describable relationship between the vegetation index NDVI and the fractional vegetation cover Fr; and (c) the deep-layer (root zone) soil water content exerts a relatively minor influence on the surface radiant temperature.

These assumptions cited above are justified by the following arguments. (a) As Price (1990) affirms, the temperature spread at low NDVI is a result of varying

degrees of surface wetness, therefore it seems reasonable to suppose that the limiting condition of the warm edge is a result of a lower limit in surface soil water content, that is, a mixture of plants and/or wholly dry surface. For this reason, the warm edge is taken to define an isopleth of limiting surface soil water content ($M_0 \approx 0$) over the range of observed NDVI. (b) Transformed vegetation indices such as NDVI have been used extensively to measure vegetation amount (Lulla 1983). Linearly equating NDVI with Fr, however, leads to errors due to the nondistributive properties of the vegetation index (Price 1990). Moreover, Paltridge and Barber (1988) together with other investigators (earlier discussion on $NDVI_0$) table caution against such equations since modification in surface conditions (e.g., background soil reflectance) and/or the physiological properties of the canopy (e.g., yellowing of the leaves) result in variations in the value of NDVI. NDVI, therefore, generally does not have a direct physical relationship with vegetation cover but is taken to vary in some monotonic fashion between bare soil and full vegetation cover, the exact form of which is determined from the simulations. (c) The SVAT model simulations (Carlson and Gillies 1991) show that the radiant surface temperature is insensitive to root zone soil water content unless the vegetation undergoes severe water stress. This assumption of a weak radiant temperature response to the root zone water content, which is critical in resolving the surface moisture availability, is based not only on the SVAT model simulations but on the implication that surface radiant temperature exhibits little spatial variability at 100% vegetation cover—that is, near the vertex of the triangle in the scatterplots.

4. Coupling the SVAT model with the satellite observations

Initial model simulations endeavor to align observed surface temperatures (i.e., from the satellite observations) with the two end points ($NDVI_s$, $NDVI_0$) along the warm edge (described by the fourth-order polynomial). This extrapolation to $NDVI_0$ and the definition of $NDVI_s$ guarantees that the implied temperatures along the warm edge for bare soil and 100% vegetation are consistent with simulations for a surface moisture availability of zero. Similarly, a solution for the deep-layer soil water content follows from establishing agreement for $NDVI_s$ at an Fr of 1.0. The deep-layer soil water content remains fixed in successive simulations for each case. In any case this estimate is not very accurate or important.

Once initialized (scaled) in this way, varying fractional vegetation (e.g., at 20% increments) for constant M_0 ($M_0 = 0$) along the warm edge resolves a relationship between NDVI and Fr in terms of the surface temperature modeled for each increment of Fr. This relationship, which is described by a fourth-order polynomial, is used to equate NDVI with Fr and can be used to map out Fr for the region (Carlson et al. 1994). Likewise, simulations continue over the theoretical range of surface moisture availability (0 to 1 in steps of say 0.2) within the full range of the fractional vegetation cover (0–1.0, i.e., corresponding to 0% and 100% vegetation cover). Thus, solutions of the model are generated by cycling over the period of the day including that of the satellite overpass with equally incremented values of M_0 and Fr until the maximum ranges of M_0 (1.0) and Fr (1.0) are reached. The net result is a matrix of simulated surface temperatures (at time of satellite overpass) for NDVI (or if preferred, Fr as computed from the relationship established for this model parameter and NDVI) at differing surface moisture availability. Henceforth, Fr and NDVI are considered interchangeable and subsequent figures may be expressed in either form.

The analogous diagrams, as generated from the SVAT simulations, for the T_0–NDVI image observations, are shown in Fig. 4. The result is a distribution of surface moisture availability isopleths that form a triangular pattern, widely spaced at low NDVI and converging at high NDVI. As previously stated, the convergence of M_0 isopleths at the vertex is due to the insensitivity of the vegetation temperature to either the deep-layer or surface soil water content. A limitation obvious upon inspection is that the (truncated) vertex of the "triangle" constitutes a zone in which isopleths of M_0 converge. Thus, the upper part of the triangle is an area where the errors in the estimated soil water content will be largest.

a. Normalization of the SVAT model simulations

Each image analysis (Fig. 4) requires an extensive set of simulations with the SVAT model, while a clear-cut warm edge may not always be apparent. Clearly, this is a somewhat involved process and requires a certain degree of experience in setting the initial conditions. Moreover, the shape, location, and size of the triangles (observed and simulated) differ from one image to the next due to variability in external (surface and atmospheric) conditions. A convenient solution lies in transforming the domain of the NDVI and T_0 so that the temperature axis does not represent a quantity highly sensitive to solar forcing, and NDVI values are scaled to fractional vegetation cover (i.e., 0–1). The transformations are as follows:

$$N^* = \frac{\mathrm{NDVI} - \mathrm{NDVI}_0}{\mathrm{NDVI}_s - \mathrm{NDVI}_0} \tag{5}$$

$$T^* = \frac{T_0 - \bar{T}_{\mathrm{air}}}{\overline{\mathrm{Rnet}}}. \tag{6}$$

NDVI (already a normalized quantity) is again normalized [Eq. (5), N^*] to scale it from 0 to 1 between the values of $NDVI_0$ for bare soil and $NDVI_s$ for 100% vegetation cover. Values of NDVI that exist outside

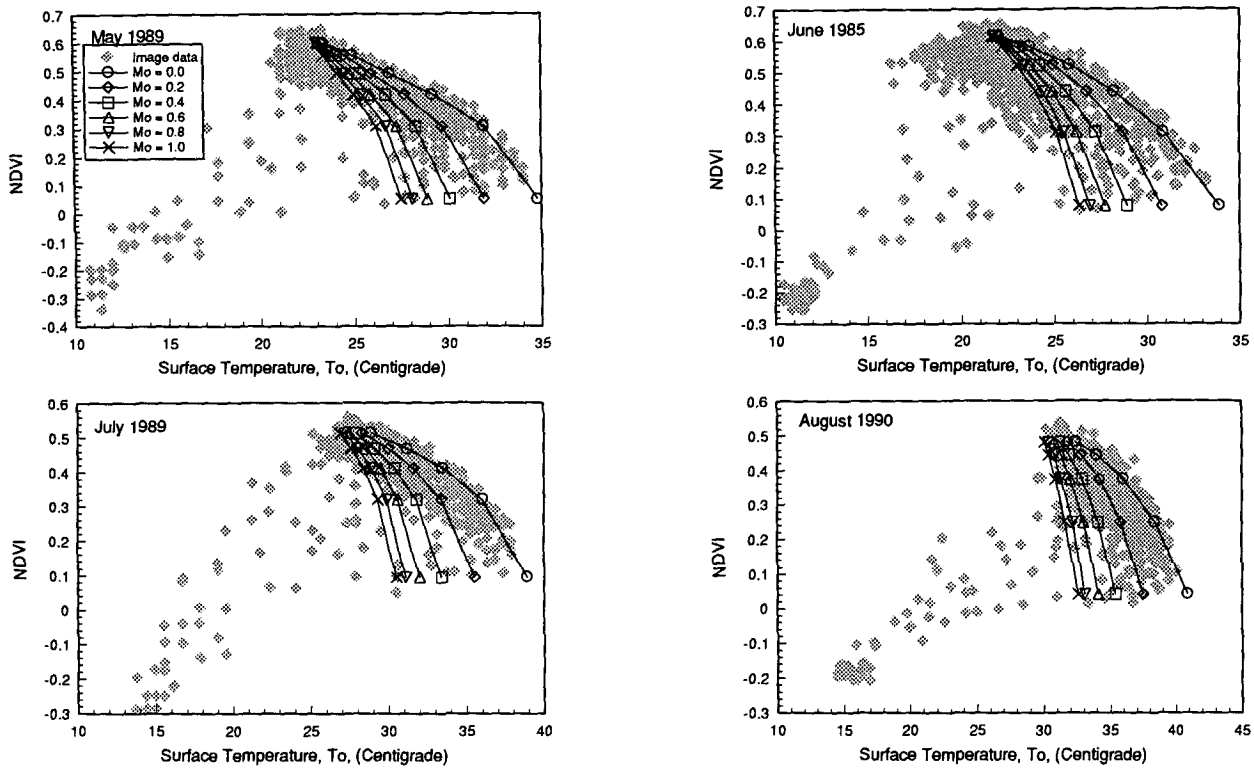

FIG. 4. NDVI versus surface radiant temperature T_0 for all four cases (as labeled) with isopleths of surface moisture availability overlaid.

these asymptotes are therefore assigned 100% or 0% cover, respectively.

The surface temperature T_0 is normalized with respect to the averaged quantities of the air temperature $\bar{T}_{air}$ at 50 m and the net radiation $\overline{\text{Rnet}}$. Note that Eq. (6) is really a transform of the Bowen ratio, assuming the sensible heat flux into the ground is small. Thus, the normalized temperature is $T^* = KB(1 + B)^{-1}$, where B is the Bowen ratio (H/L_eE), H represents the sensible heat flux, and L_eE is the latent heat flux; K represents a kind of conductance, the value of which depends upon wind speed and surface roughness. This is in agreement with Carlson and Buffum (1989), who demonstrate that the three parameters (besides soil moisture and vegetation fraction) that exert the greatest influence on the evapotranspiration are the net radiation, the surface roughness, and the wind speed.

The normalization procedure is now extended from Eq. (6) to account for the wind speed and the surface roughness. A satisfactory normalization (scaling) for T_0 is suggested by the formula

$$T^* = \frac{(T_0 - \bar{T}_{air})(V_g^{0.22}) \ln(z_0/z_r)}{\overline{\text{Rnet}}}, \qquad (7)$$

where T_0, $\bar{T}_{air}$, and $\overline{\text{Rnet}}$ are as defined earlier. Here V_g is the "surface" geostrophic wind—a value determined during the initialization of the SVAT model, z_0 the traditionally defined roughness length, and z_r a reference roughness length.

The power-law correction for the surface geostrophic wind is determined using a statistical method that seeks to align the observations for each day by minimizing the spread between each set of simulations. A visual representation of this is depicted in Fig. 5.

The scaling for the surface roughness is proposed as a dimensionless group (expressed as a logarithmic

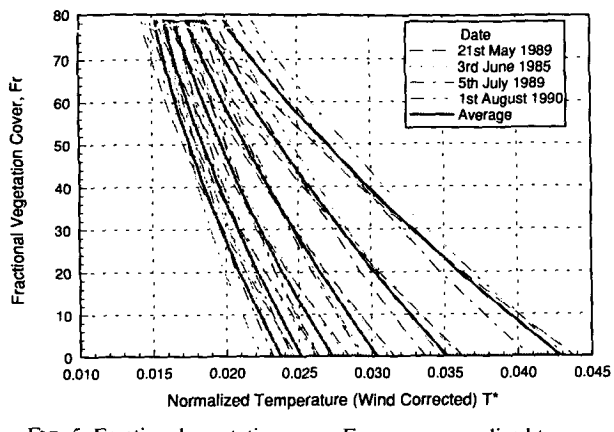

FIG. 5. Fractional vegetation cover Fr versus normalized temperature T^* for all four cases. The bold line represents the average of the individual cases. Note also that the ordinate is given in terms of Fr.

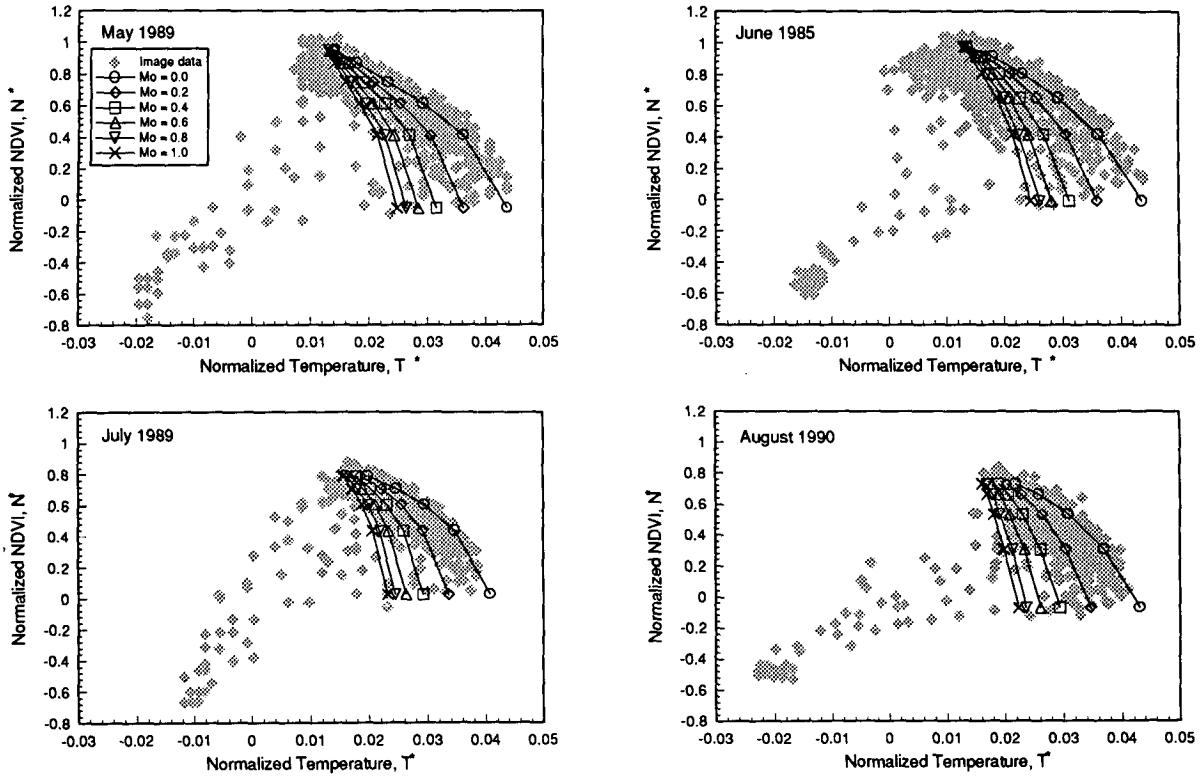

FIG. 6. Normalized NDVI N^* versus normalized temperature T^* for all four cases, with isopleths of surface moisture availability overlaid.

function) selected as relevant to the situation, but not implemented in the analyses that follow.

Figure 6 depicts the resulting triangles when normalized. The graphs clearly show that there is improvement in position and symmetry, which is to say that the triangles tend to become more congruent when normalized. Normalization therefore tends to remove differences in the radiant temperatures between images due to differences in solar elevation angle, ambient temperature, surface roughness, radiometer drift, and vegetation type. It allows one to compare images from different days and thereby to follow temporal changes in M_0 and Fr at a given point on the surface within the same triangular domain as a function of T^* and N^*, which can be computed from the remote measurements *without having to execute a full range of SVAT model calculations.*

Understandably, the normalization process does not produce perfectly congruous triangles or fit the measurements precisely, since such relatively simple transformations cannot possibly capture all the effects in the model, nor can they account for uncertainties arising from the data itself.

b. A surface fit for the surface soil water content

An important aspect of the normalization process is simplification through application of a triangle that best describes all circumstances—that is, a general form that is appropriate for most observations of T_0 and NDVI in normalized space. To this end, the isopleths of M_0 (normalized) for all four cases are averaged together to ascertain a universal fit (Fig. 7). The composite triangle (i.e., the "universal triangle") clearly fits the normalized data for August 1990 well. Therefore, the mathematical representation for the universal triangle for M_0 is given by Eq. (8):

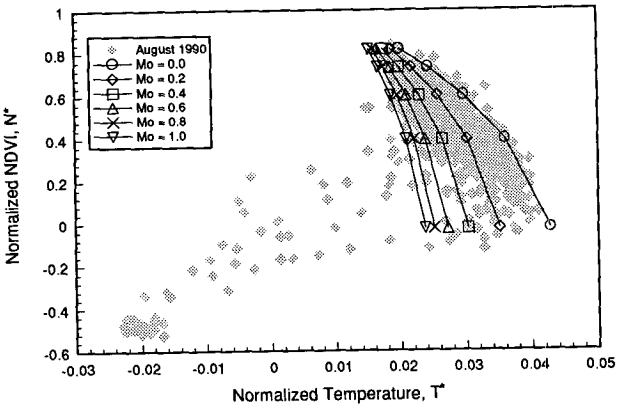

FIG. 7. Normalized NDVI N^* versus normalized temperature T^* comparing the universal triangle with observed data for a single case (1 August 1990).

$$M_0 = \sum_{p=0}^{2} \sum_{q=0}^{2} a_{pq} T^{*p} V^q, \qquad (8)$$

where V can be expressed either in terms of N^* or Fr; T^* then is wind corrected but as stated previously no account is taken for surface roughness.

c. *Maps of surface moisture availability*

The resulting maps (only August is illustrated here, Fig. 8) of M_0 reflect the apparent variability in M_0 across the terrain of the Newcastle area. An important point must be remembered when interpreting these maps. This is that the functional form of M_0, as given by Eq. (8), is only valid over the range in vegetation cover from 0% to 80%—this being the limit above which the compositing of the four triangles in Fig. 1 was found to produce unreliable results in the universal triangle.[3] This is expressed on the maps with the aid of a circle code (Fig. 8), where pixels that contain a gray circle are within the domain of the function for M_0.

In general, M_0 is observed in all four cases to have wide variability across the domain and within the region and is certainly tied to land cover, particularly noticeable in and around the built-up areas of Newcastle (exhibiting low M_0) and over the watershed area located to the northwest (upper left-hand side) where high values of M_0 are observed thereabouts. Whether there is any discernible relationship with vegetation type would be pure supposition at this stage. Further evidence that the analyses are reasonable are of a circumstantial nature. Visually correlating the maps of M_0 with a derived land use–land cover map indicates that the land to the northwest of the city is more heavily cultivated (suggesting irrigation) in comparison with that in the southwest. Interestingly, these areas exhibit a great disparity in their respective values of M_0 in a year when the previous month was climatologically very dry.

Although not shown, the overall spatial pattern of M_0 for July closely resembles that of August. However, the two other cases, May and June, have significant areas that lie outside the domain of the function for M_0 (Fig. 6), and probably reflects a variety of anomalous causes such as slope effects, shading, and standing water. Differences between the earlier and later cases may also reflect differences in agricultural and land management practices, where early crops such as winter wheat and winter barley may still be in place during May and early June.

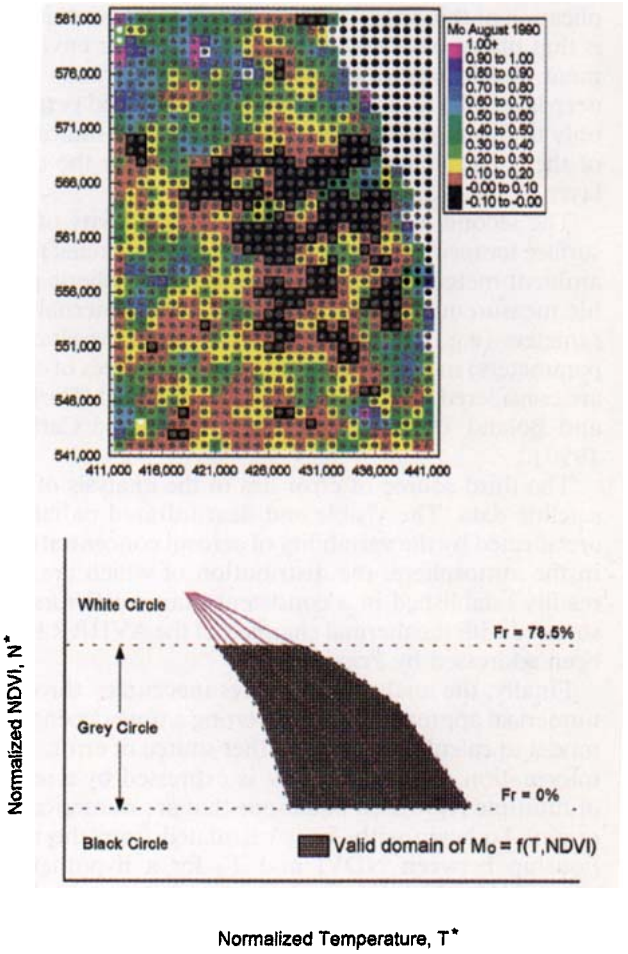

FIG. 8. Upper panel shows the map of M_0 (1 August 1990) generated as a function of normalized temperature T^* and normalized NDVI N^* for the Newcastle upon Tyne area (cf. lower panel). The lower panel is the schematic diagram showing isopleths of M_0 in N^*/T^* space showing valid domain for the derived function of M_0.

d. *Errors and uncertainties in the procedure*

In addition to errors and uncertainties arising from the three assumptions referred to earlier in the text, further independent sources of error and uncertainty exist in the method. The first constitutes possible uncertainty as a result of misinterpretation of the scatterplot where high temperatures are a result of water stress in the vegetation, particularly if that stress were to be inhomogeneous in space. We have never seen[4] any evidence of stressed vegetation in the temperature images and the evidence for this is the triangle, whose vertex implies a very small range in the observed radiant temperature over dense vegetation. This temperature tends to be close to air temperature. The im-

[3] This number was arrived at by comparing individual triangles with the composite. A large increase in the "error" with increasing Fr occurs just above 80% in Fr.

[4] This statement is based on the examination of many images at various scales and over differing types of terrain, including semiarid environments.

plication of this observation (which is counterintuitive) is that plants have evolved to optimize their environment and in any case, the water stress pertains to a deep layer (the root zone); whereas the method pertains only to the *surface* moisture, which is more indicative of the surface energy balance than that for the deep layer.

The second is associated with the sensitivity of the surface temperature to surface features (e.g., relief) and ambient meteorological conditions (atmospheric profile measurements), as well as to various internal parameters (e.g., albedo, surface roughness, vegetation parameters) in the SVAT model. These sources of error are considered in earlier literature in the field (Carlson and Boland 1978; Carlson 1986; Lynn and Carlson 1990).

The third source of error lies in the analysis of the satellite data. The visible and near-infrared radiances are affected by the variability of aerosol concentrations in the atmosphere, the distribution of which are not readily established in a consistent manner. Errors associated with the thermal channels of the AVHRR have been addressed by Price (1990).

Finally, the analysis introduces inaccuracy through numerical approximations. Inverting a time-dependent model to calculate M_0 is a further source of error. The information contained in M_0 is expressed by a series of multiple regression equations that are numerical in origin. To begin with, Fr is calculated from the relationship between NDVI and T_0 for a hypothetical warm edge, which is in itself an averaged quantity and is constrained between hypothetical limits ($NDVI_0$ and $NDVI_s$). The governing parameters Fr and T_0 (which define the numerical fit for M_0) output by the model are therefore subject to a certain amount of inaccuracy. Normalization is also a source of error. It can be shown (Fig. 9), however, that the residual error in forcing congruency between triangles is generally less than 20% of the full range in M_0 up to a fractional vegetation cover of about 80%.

Such limitations and interdependencies mean that it is difficult to come forward with a definitive assessment of the total error on the results of M_0. According to Raffy and Becker (1984), with proper constraints (i.e., the warm edge, $NDVI_0$ and $NDVI_s$) an accuracy in the inversion of a model of a few percent is achievable. Errors associated with the numerical operations in determining M_0 (rather than errors in measurement) average at $2\%(\mathrm{Fr}/T^*)$ to $3\%(N^*/T^*)$ with a maximum of 5%–7% constrained by vegetation cover. This suggests the method is probably capable of producing good results up to and over the range of 80% in fractional vegetation cover.

5. Conclusions

In practice, the most useful conceptualizations of the earth's land surface are those that match the informational requirements of the scientific community. The soil moisture availability is just such a concept, not only because this variable is a significant environmental descriptor that integrates much of the hydrology of the land surface, but also because the pattern of the moisture availability is an influential element in hydrological and meteorological processes.

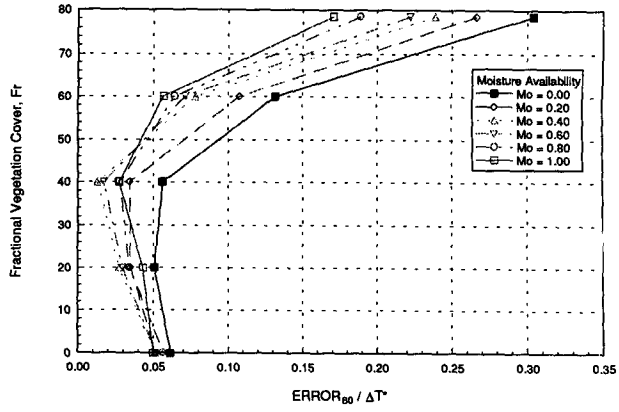

FIG. 9. Fractional vegetation cover versus an uncertainty index comprising the maximum error of estimate of the normalization procedure divided by the range of normalized temperature in each class of vegetation fraction.

Despite past and recent reservations such as those of Monteith (1981) and Hall et al. (1992) that concern the use and capability of satellite measurements in the thermal IR to infer soil moisture in the presence of vegetation, the information contained in the slope of the vegetation index–surface temperature relationship permits the resolution of this variable over partially vegetated surfaces. Unlike previous investigations, a *surface* moisture availability M_0 is determined by coupling thermal infrared temperatures along with a vegetation index, instead of a bulk moisture availability, which is obtained solely from the temperature measurements. The approach is general in that the inversion technique does not depend on specific parameterization; comparable results are obtainable with other models of similar type.

As with most boundary layer observations there exists the possibility that with the proper choice of scaling, the observed data can be applied everywhere all of the time in any situation under study. This is the idea behind the normalization of the satellite data using the four cases in question. The importance of the normalization of the data is that the triangular nomogram is less subject to variations in external forcing (solar and wind) and surface roughness and so allows the direct comparison of images made under different conditions. A universal triangle fits the Newcastle data well.

A potentially troublesome issue is in choosing a single set of parameters, the warm edge and upper and lower bounds for NDVI and a single root zone soil

water content, for an entire image. For example, $NDVI_0$ may differ for urban and rural areas. Although we have found no evidence of multiple warm edges, we cannot be sure that the warm edge constitutes an isopleth of lower limiting surface soil water content without further verification. The SVAT model, however, imposes constraints that are consistent with a theoretical limit for absolutely dry soil. So far, the possible ambiguities in the scheme where we imply a single value for these parameters over a particular surface area has not been resolved. What the scheme does resolve, however, is some of variance in the data associated with a variable vegetation cover, which strongly modulates the surface radiant temperature. A unique attribute of the moisture availability values, symbolic or otherwise, is that they pertain only to the soil surface. Thus, the method is able to separate surface from root zone moisture, although only a single value for the latter can be obtained for an image. Essential in resolving the signal due to surface soil water content from that associated with the root zone water content is the fact that the surface radiant temperature is relatively insensitive to changes in the deeper-layer soil water content, at least until the vegetation begins to experience significant water stress.

An important aspect of land surface process parameterization in mesoscale models is the initialization of the surface water content (Smith et al. 1994) so that parameterization schemes for evaporation supply the correct partition of energy and provide an accurate description of the atmospheric response. For example, Lanicci et al. (1987) show the consequence of soil moisture on the formation of mesoscale flows that intensified a front that spawned a tornadic storm. With this in mind, the universal triangle may prove to be an objective (Capehart and Carlson 1994a) means for introducing remote measurements of soil water content into the initialization of regional-scale atmospheric and hydrology models.

The universal-triangle concept by itself would generally have utility only in very short-term forecasts or theoretical studies; however, in conjunction with a long-term soil hydrology model (Capehart and Carlson 1994b), which in itself is highly dependent upon initial soil moisture conditions, remotely derived estimates (from the universal triangle) would tend to check values derived from the predicted quantities of such soil hydrology models (Ottlé et al. 1989). This would be a significant step in the development of a soil moisture parameterization scheme in operational meteorology.

Acknowledgments. We would like to thank David Ripley for his assistance in the preparation of the diagrams. The research could not have been conducted without the financial assistance of the following funding agencies. The Scientific and Engineering Council of Great Britain (SERC Contract GR/E/20080), the National Aeronautics and Space Administration (NASA Contract NAGW-2647), the United States Department of the Interior (USDI-USGS Contract 14-08-0001-G1889), and the Remote Sensing Laboratory of the USDA/ARS in Beltsville, Maryland (as part of a cooperative research agreement, Contract 58-32U4-8-27).

REFERENCES

Asrar, G., M. Fuchs, E. T. Kanemasu, and J. L. Hatfield, 1984: Estimating absorbed photosynthetic radiation and leaf area index from spectral reflectance in wheat. *Agron. J.*, **76**, 300–306.

Avissar, R., and R. A. Pielke, 1989: A parameterization of heterogeneous land surfaces for atmospheric numerical models and its impact on regional meteorology. *Mon. Wea. Rev.*, **117**, 2113–2136.

Baret, F., S. Jacquemoud, and J. F. Hanocq, 1993: The soil line concept in remote sensing. *Remote Sens. Rev.*, **7**, 65–82.

Benjamin, S. G., and T. N. Carlson, 1986: Some effects of surface heating and topography on the regional severe storm environment. Part I: Three-dimensional simulations. *Mon. Wea. Rev.*, **114**, 307–329.

Capehart, W. J., and T. N. Carlson, 1994a: Estimation of surface moisture availability using a hydrologic budget model aided by surface satellite observations and a Soil-Vegetation-Atmosphere-Transfer Scheme (SVAT). *21st Conf. on Agricultural and Forest Meteorology,* San Diego, CA, Amer. Meteor. Soc., J39–J42.

——, and ——, 1994b: Estimating near-surface soil moisture availability using a meteorologically driven soil water profile model. *J. Hydrol.*, **160**, 1–20.

Carlson, T. N., 1986: Regional-scale estimates of surface moisture availability and thermal inertia using remote thermal measurements. *Remote Sens. Rev.*, **1**, 197–247.

——, and F. E. Boland, 1978: Analysis of urban–rural canopy using a surface heat flux/temperature model. *J. Appl. Meteor.*, **17**, 998–1013.

——, and M. J. Buffum, 1989: On estimating total daily evapotranspiration from remote surface temperature measurements. *Remote Sens. Environ.*, **29**, 197–207.

——, and R. R. Gillies, 1991: Remote sensing of soil moisture over vegetation. Sensitivity and limitations of the infrared temperature method. *20th Conf. Agricultural and Forest Meteorology,* Salt Lake City, UT, Amer. Meteor. Soc., 214–217.

——, J. K. Dodd, S. G. Benjamin, and J. N. Cooper, 1981: Satellite estimation of the surface energy balance, moisture availability and thermal inertia. *J. Appl. Meteor.*, **20**, 67–87.

——, E. M. Perry, and T. J. Schmugge, 1990: Remote estimation of soil moisture availability and fractional vegetation cover for agricultural fields. *Agric. For. Meteor.*, **52**, 45–69.

——, R. R. Gillies, and E. Perry, 1994: A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover. *Remote Sens. Rev.*, **9**, 161–173.

Diak, G., 1990: Evaluation of heat flux, moisture flux and aerodynamic roughness at the land surface from knowledge of the PBL height and satellite-derived skin temperatures. *Agric. For. Meteor.*, **52**, 181–198.

Gutman, G., 1987: The derivation of vegetation indices from AVHRR data. *Int. J. Remote Sens.*, **8**, 1235–1243.

Hall, G. H., K. F. Huemmrich, S. J. Goetz, P. J. Sellers, and J. E. Nickeson, 1992: Satellite remote sensing of surface energy balance: Success, failures and unresolved issues in FIFE. *J. Geophys. Res.*, **97**, 19 061–19 089.

Huete, A. R., 1988: A soil adjusted vegetation index (SAVI). *Remote Sens. Environ.*, **25**, 295–309.

Jackson, R. D., 1983: Spectral indices in *N*-space. *Remote Sens. Environ.*, **13**, 409–421.

Jasinski, M. F., and P. S. Eagleson, 1990: Estimation of subpixel vegetation cover using red-infrared scattergrams. *IEEE Trans. Geosci. Remote Sens.*, **28**, 253–267.

Lanicci, J. M., T. N. Carlson, and T. T. Warner, 1987: Sensitivity of the Great Plains severe storms environment to soil moisture distribution. *Mon. Wea. Rev.*, **115,** 2260–2673.

Lauritson, L., G. J. Nelson, and F. W. Porto, 1988: Data extraction and calibration of Tiros-N/NOAA radiometers. NOAA, NESS. NOAA Tech. Memo. NESS 107, 58 pp.

Lulla, K., 1983: The Landsat satellites and selected aspects of physical geography. *Prog. Phys. Geogr.*, **7,** 1–45.

Lynn, B., and T. N. Carlson, 1990: A stomatal resistance model illustrating plant versus external control of transpiration. *Agric. For. Meteor.*, **52,** 5–43.

Mahfouf, J. F., E. Richard, and P. Mascart, 1987: The influence of soil and vegetation on the development of mesoscale circulations. *J. Climate Appl. Meteor.*, **26,** 1483–1493.

Markham, B. L., and J. L. Barker, 1986: Landsat MSS and TM post-calibration dynamic ranges, exoatmospheric reflectances and at-satellite temperatures. EOSAT Landsat Tech. Note, 1, 3–8.

Monteith, J. L., 1981: Evaporation and surface temperature. *Quart. J. Roy. Meteor. Soc.*, **107,** 1–27.

Nemani, R. R., and S. W. Running, 1989: Estimation of regional surface resistance to evapotranspiration from NDVI and thermal-IR AVHRR data. *J. Appl. Meteor.*, **28,** 276–284.

——, L. Pierce, S. Running, and S. Goward, 1993: Developing satellite-derived estimates of surface moisture status. *J. Appl. Meteor.*, **32,** 548–557.

Ottlé, C., D. Vidal-Madjar, and G. Girard, 1989: Remote sensing applications to hydrological modeling. *J. Hydrol.*, **105,** 369–384.

Paltridge, G., and J. Barber, 1988: Monitoring grassland dryness and fire potential in Australia with NOAA/AVHRR data. *Remote Sens. Environ.*, **25,** 381–394.

Pearson, R. L., and L. D. Miller, 1972: Remote mapping of standing crop biomass for estimation of the productivity of the short-grass prairie, Pawnee National Grasslands, Colorado. *Proc. 8th Int. Symp. on Remote Sensing of Environment,* Ann Arbor, MI, ERIM, 1357–1381.

Pielke, R. A., 1984: *Mesoscale Meteorological Modeling.* Academic Press, 612 pp.

Pinty, B., and M. M. Verstraete, 1992: GEMI: A non-linear index to monitor global vegetation from satellites. *Vegetatio,* **101,** 15–20.

Pinty, J. P., P. Mascart, E. Richard, and R. Rosset, 1989: An investigation of mesoscale flows induced by vegetation inhomogeneities using an evapotranspiration model calibrated against HAPEX-MOBILHY data. *J. Appl. Meteor.*, **28,** 976–992.

Price, J. C., 1983: Estimating surface temperatures from satellite thermal infrared data—A simple formulation for the atmospheric effect. *Remote Sens. Environ.*, **13,** 353–361.

——, 1987: Calibration of satellite radiometers and the comparison of vegetation indices. *Remote Sens. Environ.*, **21,** 15–27.

——, 1990: Using spatial context in satellite data to infer regional scale evapotranspiration. *IEEE Trans. Geosci. Remote Sens.*, **28,** 940–948.

Raffey, M., and F. Becker, 1984: Inverse problems and solution for remote sensing infrared bands. *J. Geophys. Res.*, **90,** 5809–5819.

Soares, J. V., R. Bernard, O. Taconet, D. Vidal-Madjar, and A. Weill, 1988: Estimation of bare soil evapotranspiration from microwave measurements. *J. Hydrol.*, **99,** 281–296.

——, S. Jiancheng, J. Vanzyl, and T. Engman, 1992: Estimation of bare soil evaporation using multifrequency airborne SAR. *Int. Geoscience and Remote Sensing Symp.*, Houston, TX, 1747–1749.

Smith, C. B., M. N. Lakhtakia, W. J. Capehart, and T. N. Carlson, 1994: Initialization of soil-water content in regional-scale atmospheric prediction models. *Bull. Amer. Meteor. Soc.*, **75,** 585–593.

Taconet, O., R. Bernard, and D. Vidal-Madjar, 1986: Evapotranspiration over an agricultural region using a surface flux/temperature model based on NOAA AVHRR data. *J. Climate Appl. Meteor.*, **25,** 284–307.

Tucker, C. J., 1986: Maximum normalized difference vegetation index images for sub-Saharan Africa for 1983–1985. *Int. J. Remote Sens.*, **7,** 1383–1384.