INT. J. REMOTE SENSING, 2001, VOL. 22, NO. 7, 1305–1334

Estimation of urban vegetation abundance by spectral mixture analysis

C. SMALL

Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964, USA; e-mail: small@LDEO.columbia.edu

(*Received 26 May 1999; in final form 22 February 2000*)

Abstract. The spatio-temporal distribution of vegetation is a fundamental component of the urban environment that can be quantified using multispectral imagery. However, spectral heterogeneity at scales comparable to sensor resolution limits the utility of conventional hard classification methods with multispectral reflectance data in urban areas. Spectral mixture models may provide a physically based solution to the problem of spectral heterogeneity. The objective of this study is to examine the applicability of linear spectral mixture models to the estimation of urban vegetation abundance using Landsat Thematic Mapper (TM) data. The inherent dimensionality of TM imagery of the New York City area suggests that urban reflectance measurements may be described by linear mixing between high albedo, low albedo and vegetative endmembers. A three-component linear mixing model provides stable, consistent estimates of vegetation fraction for both constrained and unconstrained inversions of three different endmember ensembles. Quantitative validation using vegetation abundance measurements derived from high-resolution (2 m) aerial photography shows agreement to within fractional abundances of 0.1 for vegetation fractions greater than 0.2. In contrast to the Normalised Difference Vegetation Index (NDVI), vegetation fraction estimates provide a physically based measure of areal vegetation abundance that may be more easily translated to constraints on physical quantities such as vegetative biomass and evapotranspiration.

1. Introduction

Recent estimates indicate that over 45% of the world's human population now lives in urban areas, with this figure rising to over 60% projected by 2030 (United Nations 1997). The global rate of urbanization is expected to continue to accelerate in the near future with the emergence of very large urban agglomerations in developing countries (Berry 1990, United Nations 1997). Even if developing countries follow the course of post-industrial urban dispersion to the suburbs, the continuing localization of populations from rural to urban/suburban conurbations will result in increasing numbers of people living in built environments. As the size and number of urban agglomerations increases, so does the relative importance of the urban environment to the global population. Monitoring spatio-temporal changes in large urban/suburban areas will therefore become increasingly important as the number and proportion of urban residents continue to increase.

The spatio-temporal distribution of vegetation is a fundamental component of the urban/suburban environment. Vegetation influences urban environmental conditions and energy fluxes by selective reflection and absorption of solar radiation (e.g.

International Journal of Remote Sensing
ISSN 0143-1161 print/ISSN 1366-5901 online
http://www.tandf.co.uk/journals

Landsberg 1981, Oke 1982, Goward *et al.* 1985, Roth *et al.* 1989, Gallo *et al.* 1993, Nichol 1996) and by modulation of evapotranspiration (e.g. Price 1990, Carlson *et al.* 1994, Gillies *et al.* 1997, Owen *et al.* 1998, Nemani and Running 1989). The presence and abundance of vegetation in urban areas may also influence air quality and human health (Wagrowski and Hites 1997). Conversely, urban vegetation experiences both short- and long-term phenological changes which may be sensitive to subtle changes in environmental conditions. Changes in the built component of the urban environment are generally documented at various levels of detail but phenological changes in urban vegetation are not under direct human control and are not generally monitored.

The synoptic view of urban landcover provided by satellite and airborne sensors is an important complement to *in situ* measurements of physical, environmental and socioeconomic variables in urban settings. Forster (1983) provides a thorough summary of the evolution of urban remote sensing and introduces a methodology with which some socioeconomic parameters may be predicted using reflectance-based estimates of landcover classes. Compared to agricultural areas and sparsely populated regions, however, application of remotely sensed observations to studies of the urban environment has been rather limited. In part this is because accurate identification of most built components of the urban environment requires finer spatial resolution than is offered by sensors on operational satellites such as Landsat and SPOT (Système Pour l'Observation de la Terre). The 30 to 50 m spatial resolution of the Landsat TM sensor (Markham 1985, Wilson 1988) is comparable to the characteristic scale of urban landcover (Welch 1982, Woodcock and Strahler 1987) and is generally too coarse for identification of individual structures. While this resolution has limited the use of Landsat data for studies of the built urban environment, it may be sufficient to detect significant spatial and temporal variations in urban vegetation.

The objective of this study is to determine the extent to which Landsat TM data can be used to quantify vegetation abundance and distribution in urban and suburban areas. In addition to the synoptic view of large urban areas, Landsat's 16-day revisit period allows monitoring of seasonal to interannual changes in urban and suburban vegetation. If the spectral resolution of Landsat TM imagery is sufficient to detect intra-urban and suburban variations in vegetation abundance, it may be possible to combine the spatial resolution of Landsat with the higher temporal and spectral resolution of sensors such as MODIS (Moderate Resolution Imaging Spectrometer) to monitor spatio-temporal variations in urban/suburban vegetation. It may also be possible to investigate temporal changes in urban/suburban vegetation abundance that have occurred during the past 18 years that the TM instruments have been operational.

2. Overview of the analysis

This analysis is based on the assumption that varying amounts and distributions of vegetation within the urban mosaic make a distinct contribution to the net reflectance observed by the Landsat TM sensor and that this contribution can be estimated by inversion of spectral mixture models of the observed reflectance. Section 3 introduces the problem of mixed pixels in the context of urban landcover and spectral mixing. Section 4 summarizes the theory of spectral mixing and defines the linear unmixing problem. The central questions of image dimensionality (§5) and spectral endmembers (§6) are then addressed for the case of urban vegetation estimation. The unmixing model previously defined is then inverted to produce vegetation

fraction estimates for different ensembles of endmembers with different constraints. The convergence of the solutions is quantified to determine the stability of the inversion and an analysis of the agreement between the estimated fractions and the observed reflectances is conducted to verify the mathematical validity of the three-component linear mixing model.

Mathematical validity does not guarantee that fraction estimates are accurate so the second part of the study focuses on validation of the estimates. A methodology is developed for calculating areal vegetation fractions from high resolution aerial photography (§8) and for simulating the TM sensor's response to the measured fractions. The effects of spatial scale and uncertainty in the geolocation of the satellite measurements are then demonstrated for different urban vegetation distributions and a validation methodology is developed (§9). Finally, a quantitative validation of the unmixed TM vegetation fraction estimates is presented for a wide range of vegetation distributions in central Manhattan (§10). The study is concluded with a discussion of the caveats of the method and the implications of the results.

3. Spectral classification of urban landcover

Urban areas are generally recognized in remotely sensed imagery by their geometric and textural characteristics. Spectral characteristics of urban landcover are less diagnostic than those of the rural periphery and unpopulated areas such as deserts and forests. There are significant differences between the spectral reflectance of urban surfaces and natural rock and soil surfaces but these differences may be difficult to detect with the limited spectral resolution provided by broadband sensors such as Landsat TM. In a study of the Washington DC metro area, Ormsby (1992) distinguished nine urban landcover classes with significant spectral separability. While these classes may be distinguishable when they occur in homogeneous regions larger than the Ground Instantaneous Field Of View (GIFOV) of the sensor, the characteristic scale at which urban landcover changes is often comparable to, or smaller than, the GIFOV of the TM sensor. Urban areas are therefore generally characterized by spectral heterogeneity at scales approaching TM pixel resolution.

The characteristic spatial scale and the spectral variability of urban landcover poses serious problems for traditional image classification algorithms. Most 'hard' classification algorithms attempt to reduce the information contained in multiband spectral reflectance imagery to a single thematic map by assigning each multiband pixel to one of a limited number of classes. The thematic class to which a pixel is assigned is generally determined by some measure of the similarity of the spectral reflectance vector of that pixel to the characteristic spectral reflectance vector of a particular class. This similarity is based on some measure of proximity to the region of the spectral feature space associated with that landcover class (Richards 1993, Jensen 1996). In areas where the landcover types have distinctive reflectance spectra and are relatively homogeneous at scales larger than the GIFOV of the sensor, the assignment of each pixel to one specific class can be done with reasonable accuracy. In areas where the reflectance spectra of the landcover vary appreciably at scales comparable to, or smaller than, the GIFOV the spectral reflectance of an individual pixel will generally not resemble the reflectance of a single landcover class but rather a mixture of the reflectances of two or more classes present within the GIFOV. Hard classifiers assign these mixed pixels to the class to which they bear the greatest statistical resemblance.

Because hard classifiers assign each pixel to one and only one class, the assignment

of a mixed pixel to a single homogeneous class produces inaccuracies in the resulting thematic map. For this reason, traditional hard classification of Landsat imagery rarely produces accurate results in urban areas where surface reflectances vary appreciably at scales smaller than the ~50 m GIFOV of the TM sensor. The problem of mixed pixels was recognized prior to the launch of Landsat 1 by Horowitz *et al.* (1971) and its relevance to urban remote sensing is discussed extensively by Forster (1985). A variety of approaches have since been developed to detect the presence and abundance of target materials within mixed pixels—particularly for use with imaging spectrometers in geologic remote sensing (e.g. Clark and Roush 1984, Goetz *et al.* 1985, Kruse 1988, Boardman 1989, Boardman and Kruse 1994) and vegetation mapping (e.g. Smith *et al.* 1985, Pech *et al.* 1986, Elvidge *et al.* 1993, Roberts *et al.* 1993, Wessman *et al.* 1994).

4. Linear spectral unmixing

Macroscopic combinations of homogeneous 'endmember' materials within the GIFOV often produce a composite reflectance spectrum that can be described as a linear combination of the spectra of the endmembers (Singer and McCord 1979). Nonlinearity is introduced by multiple scattering of radiation among different target materials and is generally considered to be a second-order effect that becomes dominant in the case of intimate mixtures (Clark and Lucey 1984). Reflectance modelling also suggests that multiple scattering can introduce significant nonlinearity for soil/vegetation-based targets (Borel and Gerstl 1994). If mixing between the endmember spectra is predominantly linear and the endmembers are known *a priori*, it may be possible to 'unmix' individual pixels by estimating the fraction of each endmember in the composite reflectance of a mixed pixel (Adams *et al.* 1986, Gillespie *et al.* 1990, Smith *et al.* 1990).

The linear mixing model assumes that the net spectral reflectance profile of an area within the GIFOV of the sensor can be described as a linear combination of endmember spectra as:

$$f_1 E_1(\lambda) + f_2 E_2(\lambda) + \cdots + f_n E_n(\lambda) = R(\lambda) \tag{1}$$

where $R(\lambda)$ is the observed reflectance profile, a continuous function of wavelength λ. The $E_i(\lambda)$ are the endmember spectra and the f_i are the corresponding fractions of the n endmembers contributing to the composite reflectance spectrum for the area within the GIFOV. Because the sensor actually measures radiance in a finite number of spectral bands, the surface reflectances are estimated from the measured radiances by compensating for sensor calibration and interaction with the atmosphere. As a result, the continuous reflectance profiles are represented as vectors of discrete reflectance estimates at specific wavelengths as:

$$E(\lambda) = [e_{\lambda_1}, e_{\lambda_2} \cdots e_{\lambda_n}] \tag{2}$$

$$R(\lambda) = [r_{\lambda_1}, r_{\lambda_2} \cdots r_{\lambda_n}] \tag{3}$$

where each r_{λ_i} represents a portion of the observed reflectance spectrum $R(\lambda)$, integrated over a finite spectral band with a centre wavelength λ_i and each e_{λ_i} represents the contribution to this observed reflectance from the corresponding endmember $E(\lambda)$. The continuous linear mixing model can therefore be represented in discrete form as a system of linear mixing equations in which the dimensions are

determined by the number of spectral bands and the number of endmembers:

$$f_j \mathbf{e}_{ij} = r_i \qquad i = 1,b \quad \text{and} \quad j = 1,n \tag{4}$$

where $\mathbf{e}_{ij}$ is the reflectance of the ith spectral band in the jth endmember spectrum, f_j is the fraction of the jth of the n endmember spectra and r_i is the ith observed reflectance of the b spectral bands. The system of b linear equations can be written as:

$$\mathbf{E}\mathbf{f} = \mathbf{r} \tag{5}$$

where $\mathbf{E}$ is a b by n matrix in which the columns are the n endmember spectra as represented by the b discrete bands and $\mathbf{f}$ is a vector of endmember fractions describing the observed reflectance vector $\mathbf{r}$. The objective is to solve for $\mathbf{f}$ to determine the fraction of each endmember that best describes the observed reflectance vector.

The tractability of the unmixing problem is determined by the discrete sampling of the reflectance spectra and the number of discernible endmembers that contribute to the target reflectances. In the formulation above, the number of bands must be equal to or greater than the number of endmembers or the problem will be underdetermined and an infinite number of possible solutions may exist. Conversely, if the number of endmembers is less than the number of bands, the problem is overdetermined and there may be no exact solution if the data contain measurement error. In this case, the mixture model can be modified to accommodate errors in the reflectance estimates that would corrupt an exact solution. The overdetermined linear mixing model, incorporating measurement error, is now:

$$\mathbf{r} = \mathbf{E}\mathbf{f} + \varepsilon \tag{6}$$

where ε is an error vector which must be minimized to find the fraction vector $\mathbf{f}$ which gives the best fit to the observed reflectance vector $\mathbf{r}$. Since $\varepsilon = \mathbf{r} - \mathbf{E}\mathbf{f}$, we seek to minimize:

$$\varepsilon^{\mathrm{T}}\varepsilon = (\mathbf{r} - \mathbf{E}\mathbf{f})(\mathbf{r} - \mathbf{E}\mathbf{f}) \tag{7}$$

A variety of approaches exist to solve this type of problem (e.g. Smith *et al.* 1985, 1990, Pech *et al.* 1986, Menke 1989, Settle and Drake 1993). The choice of approach is determined by the spectral characteristics of the endmembers and the nature of the noise component. In the case of uncorrelated noise, the well known least-squares solution is given by:

$$\mathbf{f} = (\mathbf{E}^{\mathrm{T}}\mathbf{E})^{-1}\mathbf{E}^{\mathrm{T}}\mathbf{r} \tag{8}$$

The case of correlated noise is discussed in detail by Settle and Drake (1993).

The result is a set of endmember fraction estimates for each pixel which together provide a fraction image for each endmember, indicating the spatial distribution and fraction of that endmember. The endmember fraction for a given pixel represents the fractional area of the pixel containing that endmember. In this study, the primary objective is the estimation of the vegetation fraction, F_{V}.

There exist a number of variations on this problem which incorporate different constraints on the basic estimation problem described above. Because all of the fractions should sum to unity, a constraint equation ($\Sigma f_j = 1$) can be incorporated into the problem as well as positivity constraints ($f_j \geqslant 0$) on the fraction estimates. The sensitivity of the fraction estimates to these constraints is discussed below. Settle

and Drake (1993) present a thorough discussion of the theory and application of this class of mixing model and propose a regularized solution that is less sensitive to noise and produces smoother fraction images. The regularized solution is not used in this study because a smoothness constraint is not consistent with the expectation of landcover heterogeneity at scales approaching pixel resolution.

In reality, the validity of the linear mixture model is determined by the ability of the sensor to resolve differences in the characteristic spectra of the target material given a limited sampling of the continuous reflectance spectra with a finite number of bands. In the case of imaging spectrometers, the spectra are intentionally oversampled relative to the narrowest absorption bands, but in the case of typical multispectral sensors such as Landsat, the spectra are generally undersampled so that materials with different reflectance spectra can yield indistinguishable reflectance vectors in the six bands that are imaged. In a sense this can simplify the mixing problem by reducing the number of possible endmembers that can be resolved. Conversely, this limits the applicability of Landsat data to mixing problems in which six or less endmembers describe most of the variance in the image and can be clearly distinguished in the spectral bands measured by the sensor. In this sense, the suitability of the linear mixing model is determined by the inherent dimensionality of the image and the presence of correlations between the observed bands. It is therefore necessary to determine the feature space topology of the image in order to determine whether or not a linear mixture model is appropriate.

5. Image dimensionality

The inherent dimensionality of a reflectance image is determined by the reflectance spectra of the targets and the way the sensor samples these spectra. The sampling is determined by both the spatial and spectral resolution of the sensor. Correlations among bands reduce the information content of the data. The presence of noise further reduces the information content of the data. If it is assumed that noise is spatially uncorrelated and that information content is related to spatially correlated variance, it is possible to obtain some indication of the inherent dimensionality and variance partitioning of multidimensional datasets by investigation of their eigenvalue distribution (Strang 1986, Preisendorfer 1987). Principal component transformations are routinely used to minimize the influence of band to band correlations and to investigate variance partitioning of multiband images (e.g. Richards 1993) but unequal scaling of reflectance bands can result in the noise content of one band having larger variance than the signal in another band. For this reason it is often preferable to partition variance on the basis of signal to noise ratio or spatial autocorrelation (Lee *et al.* 1990).

The Maximum Noise Fraction (MNF) transformation, proposed by Green *et al.* (1988), seeks to diagonalize the noise covariance matrix prior to the principal component transformation thereby reducing the effect of band-specific noise sources. In some cases, a noise covariance model can be determined from sensor calibration measurements or ground truth reflectance measurements of a spectrally homogeneous target. In many cases, however, these measurements are not available and it is necessary to base the noise model on the sample covariance matrix as proposed by Green *et al.* (1988). This method assumes that noise is spatially uncorrelated and may therefore be sub-optimal in the case of urban imagery where the actual reflectance varies appreciably at pixel resolution. By partitioning some of the information content to the higher order components, the MNF transformation may

underestimate the dimensionality of urban imagery. This is considered further in the error analysis below (see §8).

The six reflected bands of a Landsat TM image (scene ID LT5014032009620210) showing Manhattan island and surrounding parts of New York City and New Jersey are shown in figure 1 along with the resulting MNF transformation. Image digital numbers (DNs) were converted to normalized exoatmospheric reflectances using prelaunch calibration data for the Landsat 5 sensor (Markham and Barker 1987) but no atmospheric corrections were applied. The image was acquired at 9:18 a.m. local time on 20 July, 1996. Varying degrees of correlation among the six reflected bands suggest that the inherent dimensionality of this image may be less than the six spectral dimensions acquired. Strong correlations are apparent between the three visible bands (1, 2, 3). The first two MNF bands show strong spatially coherent contrasts illustrating spectral distinction between water, clouds, shadows, vegetation and built urban areas while the higher order MNF bands show consistently diminishing spatial coherence. The eigenvalue spectrum shown in figure 2 verifies that the first two MNF bands contain most of the spatially coherent variance.

The distribution of the transformed reflectances within the three-dimensional feature space corresponding to the low-order MNF dimensions suggests that the observed reflectances might be described by a three-component mixing model (figure 2). The density shaded scattergram of MNF dimensions 1 and 2 illustrates the triangular distribution that would be expected for three component mixing while the scattergram of dimensions 1 and 3 show most of the pixels lying on or near the plane of MNF bands 1 and 2. The scattergram of bands 2 and 3 shows that the plane is slightly inclined and disperses with increasing distance from the lower corner. If the observed reflectances can actually be represented by linear combinations of a finite number of endmembers then the reflectance vectors of the pixels at the periphery of the feature space distribution would be likely candidates for endmember spectra. The three corners of the triangular distribution correspond geographically to homogeneous areas of (1) low albedo (e.g. water), (2) high albedo (e.g. clouds, sand, concrete) and (3) vegetation (e.g. grass). In truly linear mixing, all pixels would lie within a mixing space interior to a convex hull defined by the endmembers (Boardman 1993). The slight outward curvature in the distribution of pixels shown in figure 2 implies some degree of nonlinear mixing among the three endmembers. The pixels along several dropped scan lines reside in small clusters outside this planar triangular region indicating that their reflectance vectors are not described by mixtures of the three endmembers. The remaining pixels lie intermediate between the three endmembers with the majority displaced toward a mixing line between the low- and high-albedo corners. The most significant departures from the mixing plane are related to higher dimensional reflectance spectra associated with clouds and the distinction between water and other low-albedo reflectance such as shadows.

The triangular distribution in MNF space bears an obvious similarity to the well known Tasseled Cap distribution discovered by Kauth and Thomas (1976). The distributions are similar in the sense that both contain a vegetation endmember that is distinct from a mixing continuum between absorptive and reflective endmembers. The primary difference is that the 'plane of soils' frequently observed in agricultural regions is replaced here by a lower dimensional mixing line of built urban surfaces. Since most urban surfaces are impervious to moisture, the reflectance along this axis is presumably less time variable than the plane of soils.

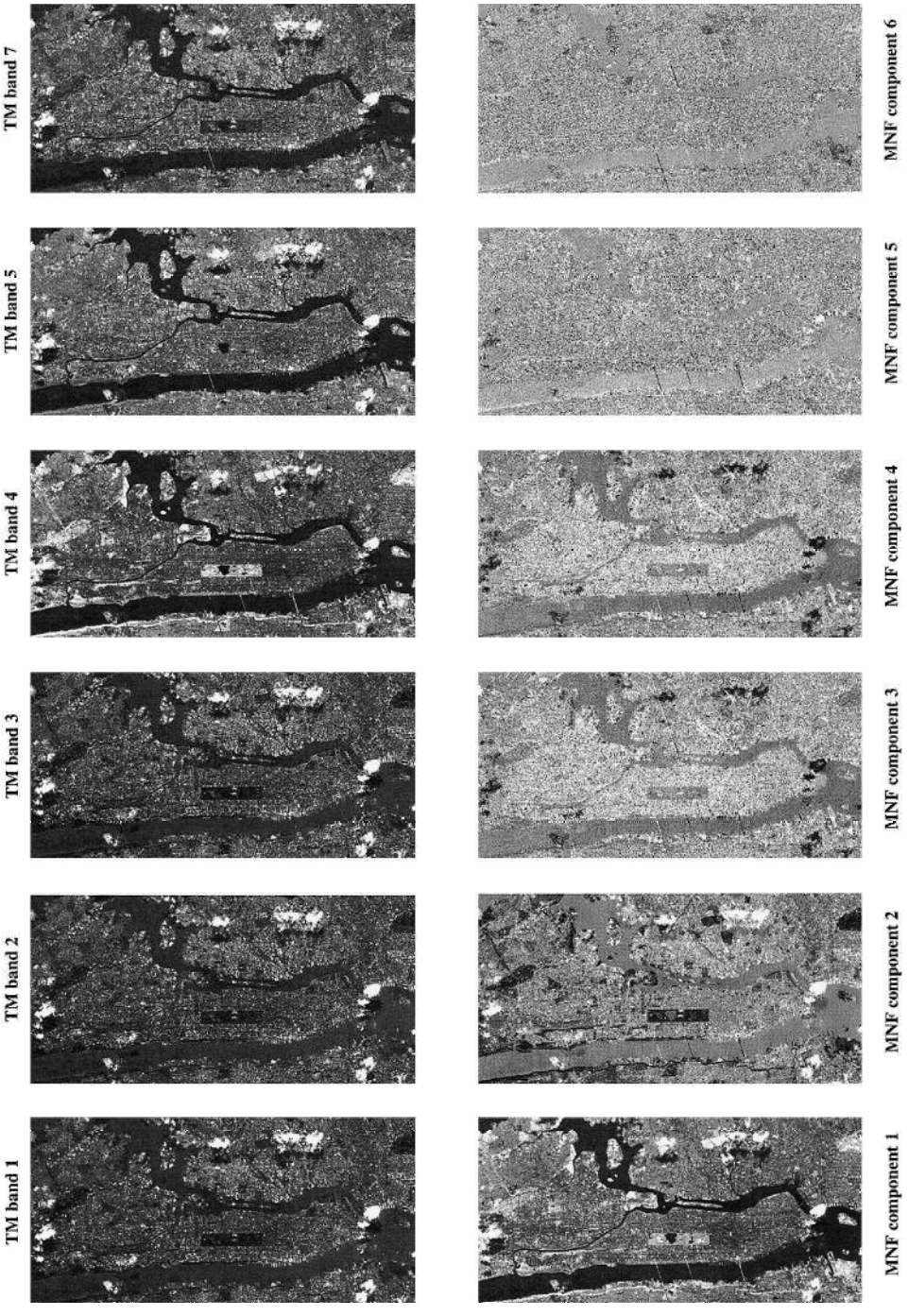


Figure 1. Dimensionality of urban multispectral reflectance. The six reflective bands of Landsat TM acquired over Manhattan Island and surrounding regions on 20 July 1996 show strong correlations, especially in the visible bands (1, 2, 3). After applying a Maximum Noise Fraction (MNF) transformation, the spatially coherent variance is contained in the low-order components (1, 2) while the spatially uncorrelated variance is relegated to the higher order components. Total information content is invariant under the transformation.

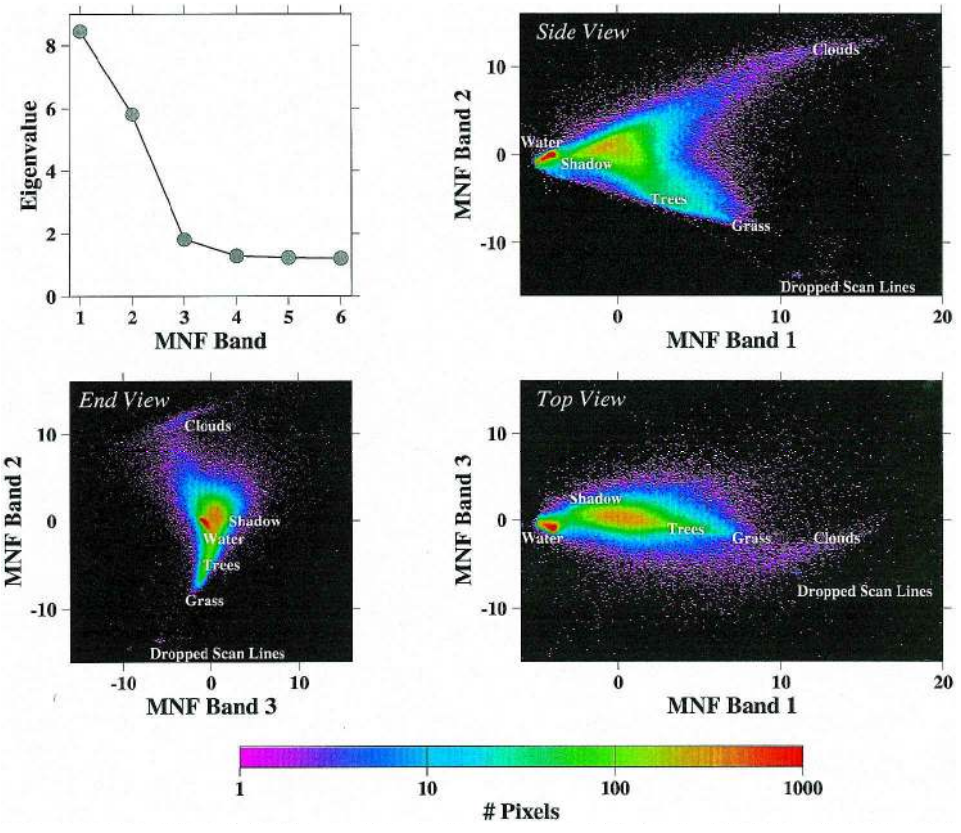


Figure 2. Transformed feature space representation of Manhattan and surrounding areas. The eigenvalue spectrum (top left) corresponding to the MNF rotated bands shown in figure 1 indicates that the first two components contain the majority of the spatially correlated variance compared to the four low-order components. Density-shaded scattergrams reveal the feature space topology of the reflectance data to have three endmembers confined primarily to the plane of MNF bands 1 and 2 (upper right). The density-shaded scattergrams show a distinction between water and shadow reflectance as well as significant offset and curvature of the cloud spectra on the upper cusp of the distribution. Urban pixels are confined primarily to the triangular region on the band 1–2 plane which is slightly inclined to the 0 axis of band 3. Note that the colour changes with the $\log_{10}$ of the number of pixels.

6. Endmember spectra

Estimation of endmember fractions from the linear mixing model is predicated on the assumption that we have some idea what spectral endmembers the mixed pixels are composed of and what materials actually constitute the spectral endmembers. The simplest mixing model is that which describes the observed reflectances with the smallest number of endmembers. The simplest mixing model that can describe the distribution in figure 2 is composed of three components associated with the apexes of the triangular distribution. If the reflectance vector of any pixel within the triangular distribution shown in figure 2 can be described as a linear combination of vectors located on the periphery of the distribution then the reflectance vectors form a convex set (Boardman 1993). Convexity has pleasing implications for the existence of solutions to linear inverse problems (Parker 1994) but it provides no

assurance of uniqueness. In other words, convexity does not guarantee that the set of endmember spectra chosen is the only set that describes the data. For this reason it is important to have some idea what the physical endmembers are. A three-component mixing model based on low albedo, high albedo and vegetation is physically consistent with the spectral characteristics that might be expected for an urban environment. Although a wide variety of target materials is present in the urban mosaic, the differences in their spectral reflectances are not necessarily resolved by the TM sensor's spectral sampling. While this would be a limitation for some applications, it may simplify the problem of vegetation detection considerably.

The success of spectral unmixing depends on the selection of the endmember components. There exist a variety of methods for choosing endmembers based on *a priori* assumptions, ground truth or laboratory reflectance measurements or the image data themselves. In many cases it is necessary to select reference endmembers to relate image endmembers to actual target materials (Roberts *et al.* 1993, Adams *et al.* 1995). In the case of urban reflectance, the relative simplicity of the feature space topology suggests that it may be possible to select endmembers from the image data. Three approaches are considered in which the endmembers are based on (1) maximally inclusive feature space selection, (2) minimally inclusive feature space selection and (3) geographic image space selection. The maximally inclusive endmembers are averages of large numbers (>2000) of reflectance profiles of all pixels near the corners of the triangular distribution in the scattergram of MNF bands 1 and 2. The minimally inclusive endmembers are averages of smaller numbers (~ 100) of reflectance profiles corresponding to the pixels nearest the apexes of the three corners of the scattergram. The geographic image space endmembers are averages of spatially homogeneous areas of cloud, water and grass corresponding to the corners of the scattergram. These three ensembles of endmembers are shown in figure 3 with continuous laboratory vegetation spectra (Clark *et al.* 1993) for comparison.

Once the image endmembers have been determined, their relationship to physically meaningful target materials must be established. The similarity of the three ensembles of endmember spectra for the low-albedo and vegetation endmembers suggests that these endmembers occupy a compact, well defined subset of spectral space. The high-albedo endmember shows much greater sensitivity to the selection method, as would be expected from the curvature and offset of the high-albedo cusp of the scattergram. The spectral endmembers determined from the transformed feature space correspond to reflectances from distinct geographic locations within the study area so the physical targets themselves can be checked in the field. The underlying assumption of endmember purity, or spectral homogeneity, can then be verified to establish the relationship between the endmembers and actual target materials. The similarity of the endmember spectra obtained from the feature space and geographic space selections supports the assumption.

The underlying question in this analysis concerns the actual physical composition of the spectral endmembers. In the New York study area, the low-albedo endmember corresponds to deep clear water and, in some cases, deep shadow. The high-albedo endmember corresponds to clouds in the minimally inclusive case with the inclusion of sand and concrete in the maximally inclusive case. In all cases the vegetation endmember corresponds to grass. This is physically reasonable because the differences observed in full-resolution vegetation spectra (Asner 1998) are not necessarily observed by the spectral sampling of the Landsat TM sensor. Dense, healthy grass has negligible internal shadowing and therefore provides the most homogeneous

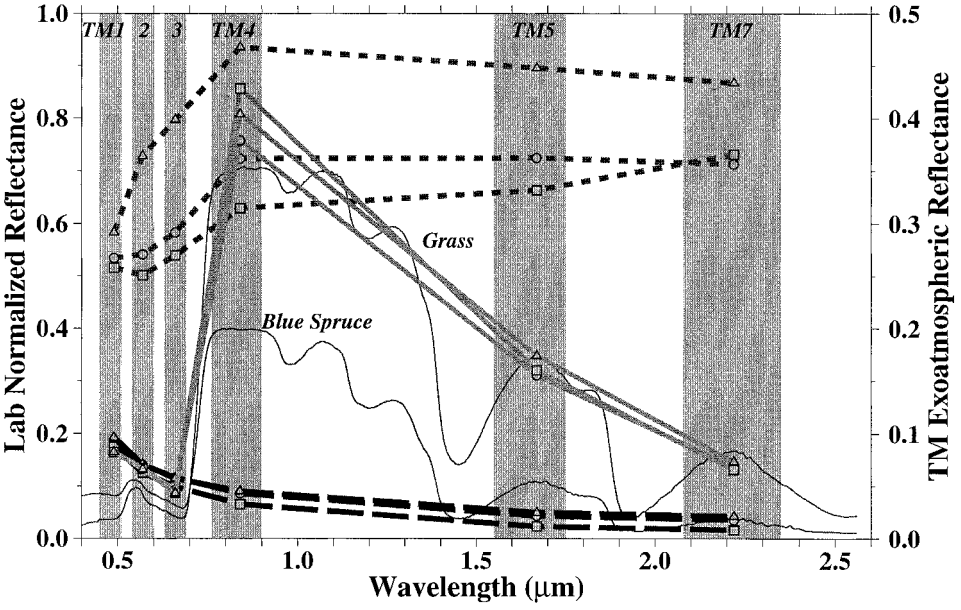


Figure 3. Endmember reflectance spectra. Discrete spectra correspond to minimally inclusive (circles) and maximally inclusive (squares) MNF space endmembers and geographically homogeneous image space (triangles) endmembers. The low-albedo (long dash) and vegetation (solid) endmembers are self-consistent and relatively invariant to the method of endmember selection. The high-albedo (short dash) spectra are more variable as a result of the spectral dissimilarity of clouds and high-albedo urban targets. TM bandwidths (grey) and continuous spectra (thin curves) derived from the United States Geological Survey spectral library are shown for comparison. Atmospheric scattering is indicated by the higher reflectance in TM band 1 for the vegetation and low-albedo endmembers.

vegetation response found in the New York metro area. Varying degrees of tree canopy closure observed in Central Park correspond to a mixing between homogeneous vegetation and varying amounts of low-albedo endmember related to shadow. Some contribution from soil is undoubtedly present, but evidently not in large enough amounts to form a distinct endmember in feature space.

7. Vegetation abundance estimates

The three-component linear mixing model was inverted for endmember fractions using the Landsat TM data shown in figure 1 with each of the three sets of endmember spectra shown in figure 3. An example of the resulting endmember fraction images is shown in figure 4. Each inversion was performed both with and without a unit sum constraint. The unit sum constraint enforces the assumption that the three endmembers completely span the space of all possible reflectance vectors. In order for the inversion to be stable, small differences in the endmember spectra should not result in large differences in the fraction estimates. The sensitivity of the solution to the choice of endmember spectra is indicated by the consistency of the endmember fraction distributions and the correlations between different solutions given in tables 1 and 2. The similarity of the distributions and the high correlations between the resulting fraction images suggest that the low-albedo fraction is most sensitive to the differences in the endmembers shown in figure 3. Enforcing the unit sum

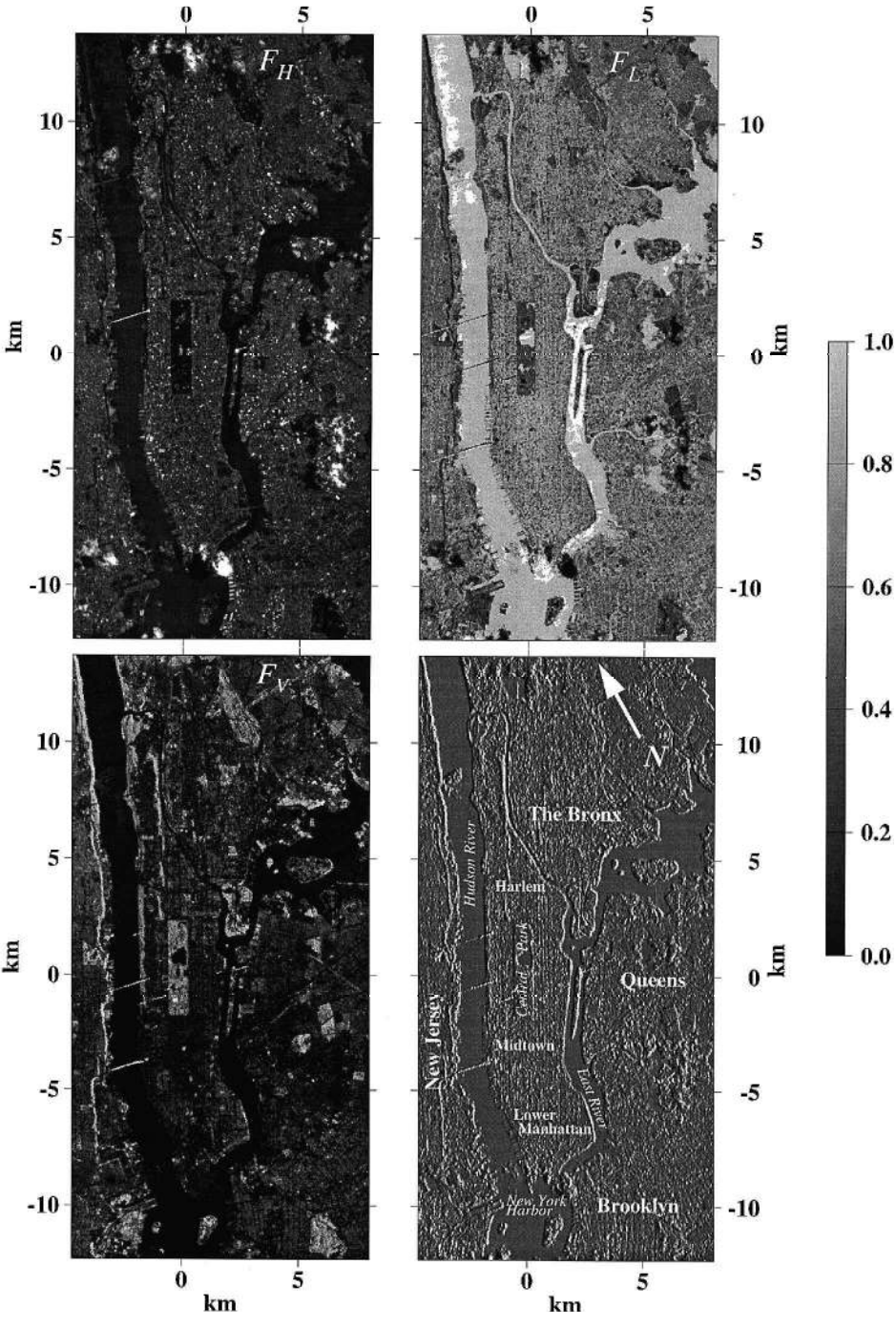


Figure 4. Three-component endmember fraction images of Manhattan and surrounding areas. Fractional abundance images result from a unit sum constrained least squares inversion of the linear mixing model using the maximally inclusive endmember spectra shown in figure 3. The parallel diagonal streaks crossing the Hudson River are dropped scan lines. The index map in the lower right shows a high pass filtered low-albedo fraction image.

Table 1. Endmember fraction and RMS misfit distributions.

	F_H		F_V		F_L		RMS	
	Mean	σ	Mean	σ	Mean	σ	Mean	σ
U_{Min}	0.20	0.17	0.15	0.19	0.49	0.39	0.007	0.007
C_{Min}	0.18	0.16	0.16	0.19	0.66	0.22	0.010	0.009
U_{Max}	0.18	0.18	0.15	0.22	0.47	0.41	0.007	0.007
C_{Max}	0.14	0.16	0.17	0.22	0.68	0.24	0.011	0.010
U_{Geo}	0.13	0.14	0.14	0.21	0.53	0.37	0.009	0.008
C_{Geo}	0.10	0.12	0.16	0.20	0.73	0.21	0.012	0.010

U = Unconstrained, F_H = High-albedo fraction, Min = Minimally inclusive endmember, C = Constrained $\Sigma F = 1$, F_L = Low-albedo fraction, Max = Maximally inclusive endmembers, F_V = Vegetation fraction, Geo = Geographic endmembers.

Table 2. Endmember fraction and RMS misfit correlation coefficients.

	U_{Min}	C_{Min}	U_{Max}	C_{Max}	U_{Geo}	C_{Geo}	
U_{Min}	1.000	0.965	0.999	0.958	0.999	0.950	F_H
C_{Min}	0.625	1.000	0.955	0.999	0.962	0.998	F_H
U_{Max}	0.999	0.638	1.000	0.947	0.999	0.938	F_H
C_{Max}	0.626	0.989	0.633	1.000	0.955	0.999	F_H
U_{Geo}	0.998	0.626	0.999	0.621	1.000	0.948	F_H
C_{Geo}	0.622	0.968	0.625	0.995	0.613	1.000	F_H
	F_L U_{Min}	F_L C_{Min}	F_L U_{Max}	F_L C_{Max}	F_L U_{Geo}	F_L C_{Geo}	
U_{Min}	1.000	0.994	0.996	0.987	0.996	0.983	F_V
C_{Min}	0.890	1.000	0.987	0.997	0.991	0.996	F_V
U_{Max}	0.965	0.901	1.000	0.987	0.999	0.982	F_V
C_{Max}	0.842	0.978	0.896	1.000	0.991	0.999	F_V
U_{Geo}	0.861	0.895	0.930	0.930	1.000	0.987	F_V
C_{Geo}	0.797	0.940	0.865	0.981	0.959	1.000	F_V
	RMS	RMS	RMS	RMS	RMS	RMS	

constraint seems to repartition the variance between the high- and low-albedo endmember fractions at the expense of slightly larger RMS misfits to the observed reflectances. This is consistent with the dispersion of the mixing space between these two endmembers seen in figure 2. In contrast, the scattergram is more planar approaching the vegetation endmember. The implication is that a three-component linear mixing problem may be ill-posed for estimation of high- or low-albedo component fractions but is well posed for vegetation component fractions. The rest of this analysis discusses the unit sum constrained solution based on the maximally inclusive endmembers, although tables 1 and 2 suggest that the results for the other solutions would not differ significantly for the vegetation fraction.

The mathematical validity of the three-component linear mixing model can be assessed by reconstructing reflectance vectors using the model endmembers and the estimated fractions. Comparing an observed reflectance vector with the corresponding vector constructed from the estimated fractions of model endmembers indicates

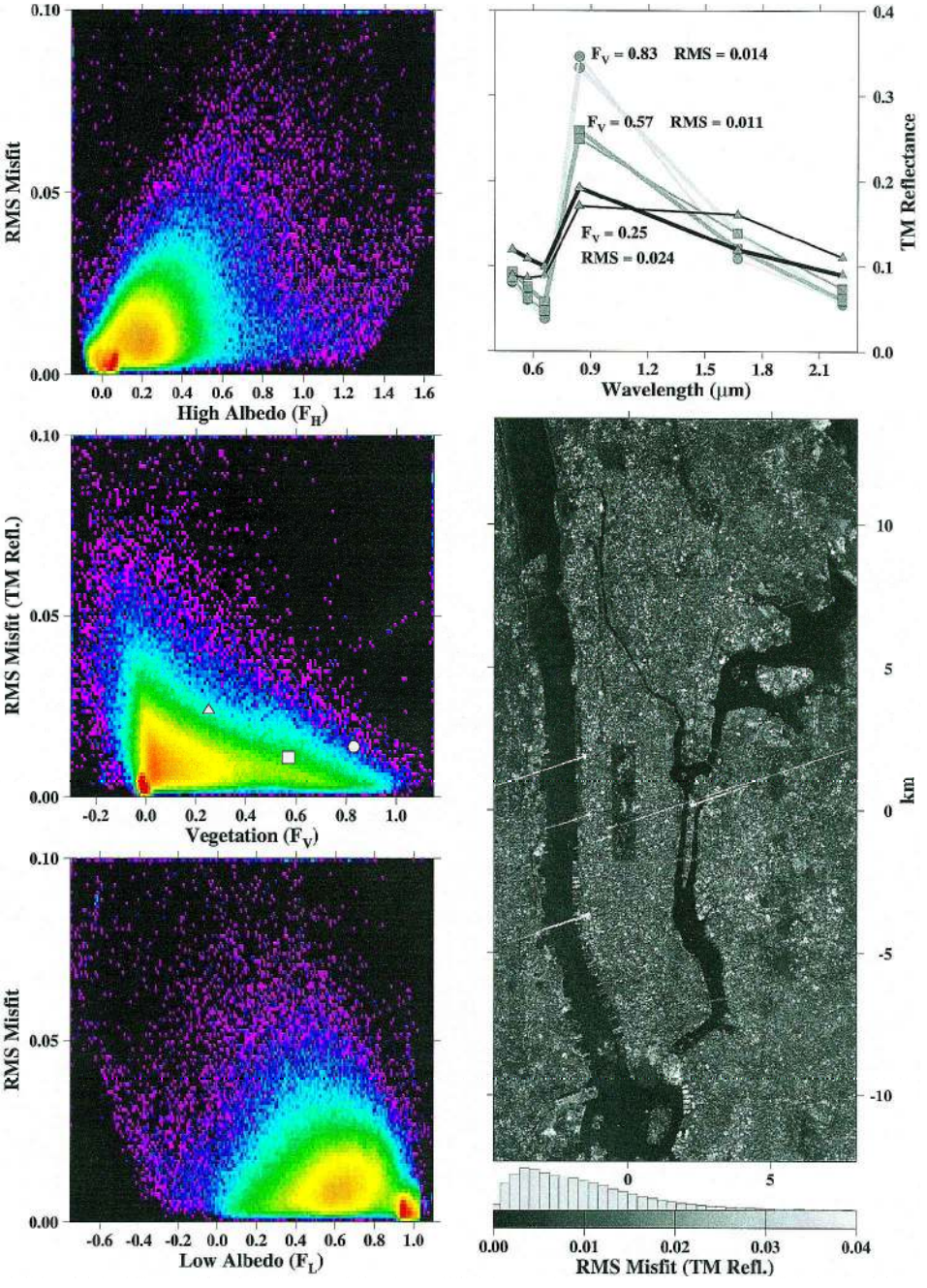
how well the linear mixing model describes the reflectance of that pixel. Example comparisons of observed and reconstructed vectors for three different vegetation fractions are shown in figure 5. The Root Mean Square (RMS) misfit between the observed and reconstructed reflectance vector gives some indication of model validity for each pixel in the image. The RMS image in figure 5 shows generally small misfit overall (95% $<$0.02 RMS) and consistently lower misfit for the low albedo (e.g. water) and vegetated areas with larger misfits for high-albedo areas, clouds and the dropped scan lines. The density-shaded scattergrams in figure 5 verify that the pixels with larger misfits are associated with larger fractions of the high-albedo endmember and intermediate fractions of the low-albedo endmember. Maximum RMS misfit diminishes monotonically with increasing vegetation fraction and is consistently low for vegetation fractions greater than 0.2. The larger misfit for high-albedo components suggests that this endmember is not accurately represented in the model but this evidently does not influence the estimation of vegetation fraction. The small misfit for pixels having a significant vegetation fraction suggests that the three-component linear mixing model may provide robust estimates of vegetation abundance in urban Landsat imagery. Low RMS misfit supports the statistical validity of the linear mixing model but does not guarantee accurate or physically meaningful results. In order to quantify the accuracy of the vegetation fraction estimates it is necessary to validate the model with independently derived ground truth estimates of the actual vegetation fraction.

8. High-resolution measurement of vegetation distribution

Measurements of the areal vegetation fraction and distribution can be obtained from high-resolution imagery in which vegetation can be unambiguously identified by texture, colour and context. If the unmixed vegetation fractions estimated from the Landsat data are similar to vegetation fractions derived from validated high-resolution imagery, the agreement will provide independent confirmation of the unmixing estimates and give some idea of the detection limit of the TM sensor. The Landsat-derived vegetation fraction estimates are validated here by comparison with high-resolution visible colour aerial photography of central Manhattan acquired by Hammon, Jensen, Wallen and Associates from an altitude of 3788 m using a calibrated, forward motion compensated, Zeiss LMK 1015 aerial survey camera. A 1:16 000 colour reproduction of this imagery is published in the *National Geographic Satellite Atlas of the World* (National Geographic Society 1998). The image was acquired nine days prior to the Landsat overpass and provides sufficient spatial and spectral resolution to allow consistent detection of vegetation at scales larger than

Figure 5. Linear mixing model misfit. RMS misfit image (lower right) indicates consistently good (95% $<$0.02 RMS) fit of the three-component mixing model to the observed reflectances. Vegetation and homogeneous low-albedo (e.g. water) pixels have low RMS misfit while high-albedo pixels and dropped scan lines have larger misfits. Examples of observed reflectance vectors (upper right, thick lines) and the corresponding reconstructed endmember reflectance (thin lines) for three different vegetation fractions show good agreement for all but the lowest vegetation fraction. Density-shaded scattergrams show larger RMS misfits for pixels with higher fractions of high-albedo or moderate fractions of low-albedo components but show consistently low misfit for vegetation fractions greater than 0.2. Symbols correspond to the three example profiles above. The colour of the scattergrams is proportional to $\log_{10}$ pixels as in figure 2.

1 m. The area imaged contains a wide variety of vegetation types including large grass covered meadows and open and closed canopy temperate forest within Central Park, numerous smaller parks with varying degrees of canopy closure, and tree-lined streets and interior courtyards in geometric arrangement. As such, it provides numerous examples of interspersed built and vegetative landcover in varying fractions



as well as large areas of complete vegetation with different degrees of internal shadowing.

Measurements of vegetation fraction are made by simulating the spatial response of the Landsat sensor to the distribution of vegetation estimated from the higher resolution aerial photograph. Vegetated pixels are assigned on the basis of visible green reflectance on the assumption that spatially contiguous visible green areas represent illuminated vegetation. This assumption is incorrect for the rivers and many of the small water bodies in Central Park because the presence of algae results in some visible green reflectance but these areas do not influence the validation. In general, however, spatially continuous areas larger than $\sim$2 m with predominantly visible green reflectance represent vegetation. Smaller isolated green areas occur infrequently and are attenuated by the spatial averaging process described below. Extensive field validation in central Manhattan further verifies the correspondence between spatially contiguous green areas and vegetation.

For the purposes of validation it is necessary to distinguish between fully illuminated vegetation, partially illuminated vegetation, shadow and non-vegetated areas. The net vegetative reflectance is determined both by the amount of boundary shadowing from adjacent buildings and internal shadowing related to canopy structure so it is important to discriminate not just between vegetated and non-vegetated areas but also between different amounts of illuminated vegetation within these areas. At scales larger than $\sim$2 m, vegetation in New York City consists primarily of deciduous trees with lesser contributions from evergreen trees and grass, so using a 2 m resolution pixel should incorporate all significant vegetated areas with a minimal number of false positives. Although there are significant differences in the green hues and saturations of different vegetation, the assumption is that at 2 m resolution the darker green areas correspond to illumination differences rather than species composition. This assumption is based on the observation that intracanopy shadows range through darker green into brighter green in illuminated canopy.

Vegetation measurements are made for 2 m square airphoto pixels by calculating the ratio of green to the sum of red and blue visible reflectance, g/(r+ b), as shown in figure 6. For equal proportions of red, green and blue, this ratio has a value of 0.5 corresponding to grey shades ranging between black and white. Most construction materials lie somewhere on or near this grey axis so the histogram corresponding to the ratio image in figure 6 shows a peak near 0.5. Partially shadowed vegetation within tree canopies appears darker green than fully illuminated vegetation such as grass. Forested areas in Central Park have g/(r+ b) ratios ranging from 0 to $\sim$1.2 with consistent modal values of $\sim$0.8. By applying a clipped linear stretch to the ratio image and assigning pixels with ratios less than 0.5 a value of zero and assigning pixels with ratios greater than 0.8 a value of 1 it is possible to generate a nearly binary image of vegetation presence or absence while allowing partially illuminated vegetation to contribute to the integrated vegetation response. This admittedly *ad hoc* method neglects the fact that illuminated vegetation spans a range of green hues but the resulting image does show the presence of vegetation remarkably well while accommodating differences in canopy structure and internal shadowing which would be expected to influence the net reflectance of fully vegetated areas.

In order to compare the vegetation fractions estimated from the aerial photograph with those obtained from the Landsat image it is necessary to account for the instrument response of the TM sensor. Convolving a 50 m $\times$ 50 m radially symmetric TM point spread function (Markham 1985) with the stretched green ratios described

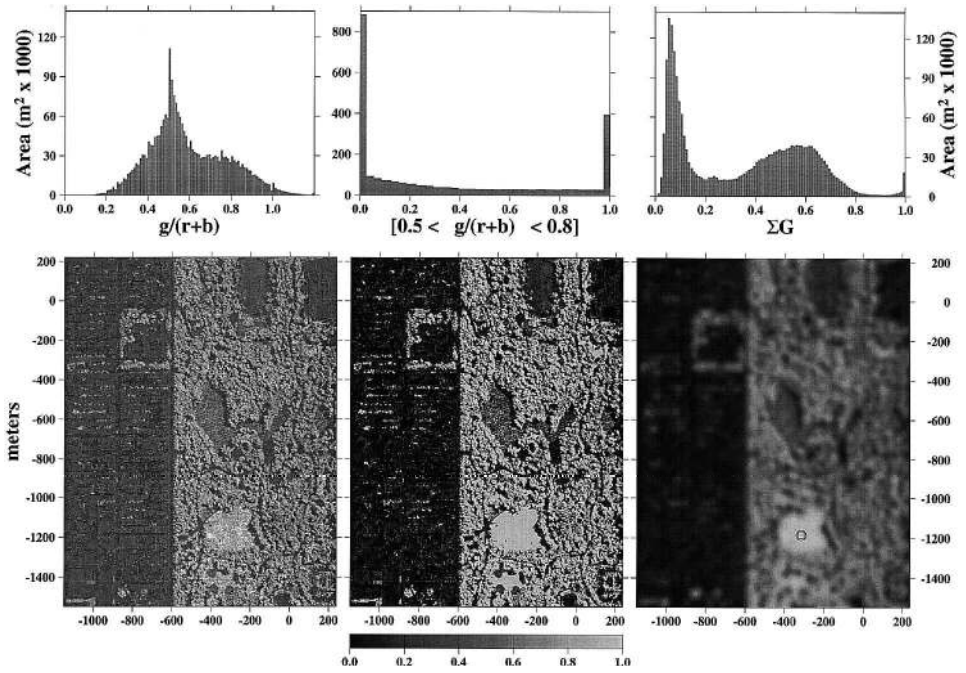


Figure 6. High-resolution estimation of vegetation abundance. A colour visible aerial photograph of Central Park and the Upper West Side of Manhattan at 2 m resolution shows the presence of vegetation by larger ratios of visible green reflectance (lighter areas) relative to visible red and blue reflectance (g/(r+ b), left image). Applying a clipped linear stretch between green ratios of 0.5 and 0.8 maps pixels along the grey axis (g/(r+ b)= 0.5) and below (g/(r+ b)<0.5) onto 0 and maps pixels containing healthy, unshadowed vegetation (e.g. grass g/(r+ b)>0.8) onto 1 (central image). Pixels with stretched values between 0 and 1 generally correspond to partially shadowed vegetation. Convolving the 2 m green ratio vegetation image with the point spread function of the Landsat 5 TM sensor yields a response weighted estimate of the vegetation fraction expected for a 28.5 m pixel centered on each 2 m airphoto pixel (right image). The size of the 50 m IFOV and 28.5 m pixel in the center of the Sheep Meadow (bottom center) relative to the scale at which the vegetation fraction changes indicates that a small uncertainty in the precise location of a TM pixel can result in significant differences in the vegetation fraction detected by the sensor. Shading is linear over the ranges indicated above each image.

above results in a response weighted estimate of the total vegetation fraction within the Landsat sensor's ~50 m GIFOV when centred on each of the 2 m airphoto pixels (figure 6). The distance-weighted measure of the vegetation fraction that would be detected by the TM sensor is given by:

$$\Sigma G(x,y) = \int_{\mathrm{IFOV}_{u,v}} R(x-u, y-v)\; G(x,y)\,\mathrm{d}u\,\mathrm{d}v \tag{9}$$

where $R(u,v)$ is the point spread function of the TM sensor within the GIFOV and $G(x,y)$ is the stretched green ratio. The resulting image appears smoother at 2 m resolution than the corresponding Landsat image with 28.5 m pixels (figure 7) and shows that the vegetation fraction within the GIFOV of the sensor can vary

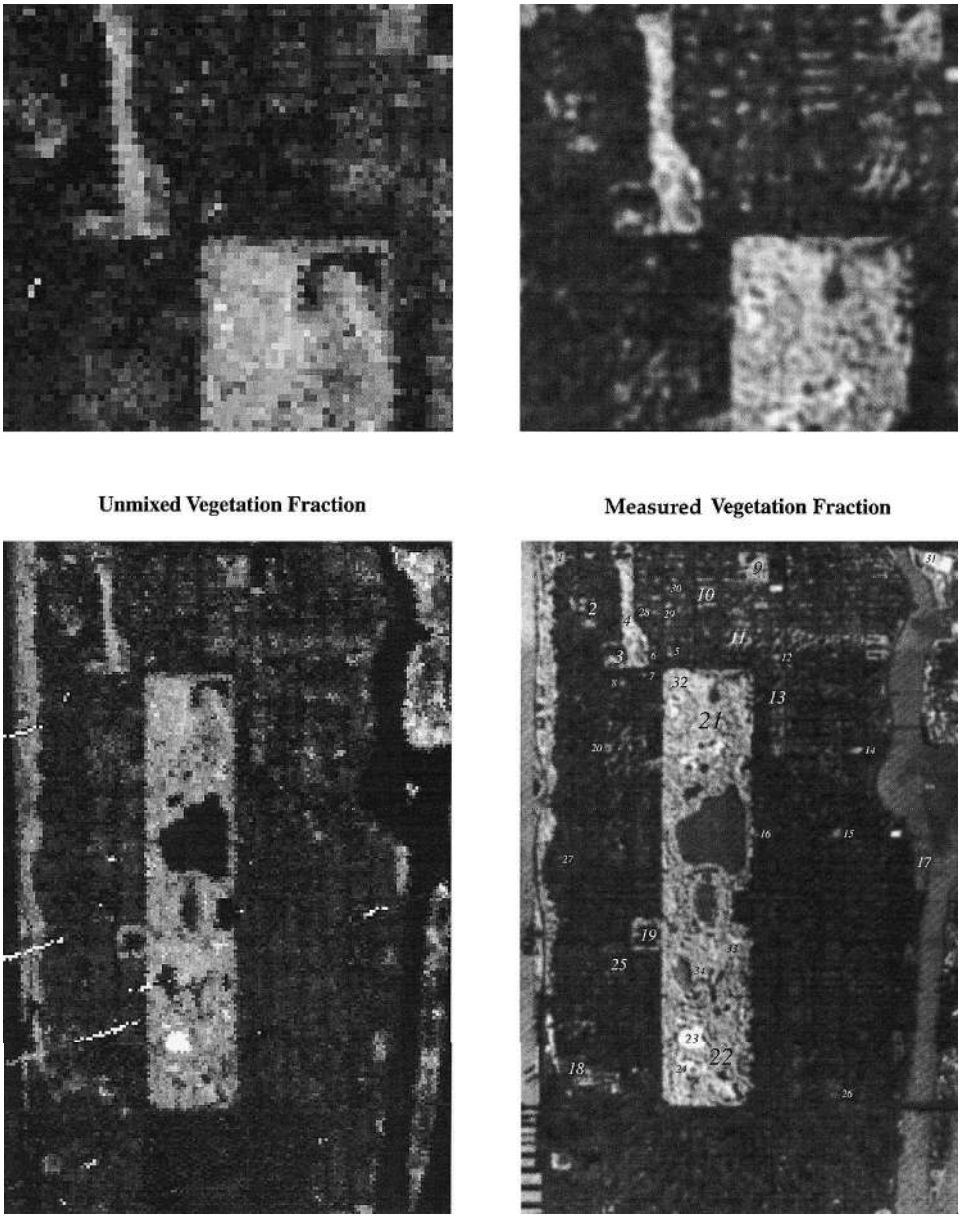


Figure 7. Comparison of measured and unmixed vegetation fractions in central Manhattan. The vegetation fraction (F_V) image derived from the aerial photograph is shown at 2 m resolution while the unmixed Landsat TM vegetation fraction image on the left is limited to 28.5 m spatial resolution. Shading is linear for vegetation fractions between 0 .0 (black) and 1.0 (white) for both images. Non-zero F_V for the Hudson and East rivers in the 2m image is a result of visible green scattering by algae and suspended particulate matter in the water. The numbers show locations of validation sites (table 3).

appreciably at scales smaller than a single 28.5 m Landsat pixel. This observation has profound implications for validation of the endmember fraction estimates.

In order to validate the vegetation fractions unmixed from the Landsat data it

is necessary to compare them with the higher resolution vegetation fractions measured from the airphoto. To compare the vegetation fraction of a single 28.5 m Landsat pixel with the measured vegetation distribution in the corresponding 50 m × 50 m GIFOV on the airphoto, it is necessary to know the precise location of the GIFOV. This presents a problem if the response-weighted 2 m vegetation fraction estimate varies significantly at scales smaller than 28.5 m because it is not generally possible to determine the location of a Landsat pixel to subpixel precision. As a result, small (<14 m) uncertainties in the location the GIFOV can result in significant differences in the vegetation fraction measured from the airphoto that is expected to correspond to the single Landsat pixel.

9. Characteristic scale and fraction estimates

The distance over which the response-weighted vegetation fraction changes is determined by the characteristic scale of the vegetation distribution and the characteristic scale of internal shadowing in the tree canopy. One measure of this characteristic scale can be obtained from the spatial autocorrelation function by the width of its central peak. This lag distance corresponds to the average distance between maxima and minima of the vegetation fraction image and provides a quantitative indication of the characteristic scale of the vegetation distribution. If the distribution is not isotropic the scale will depend on azimuth. Figure 8 shows radial autocorrelation profiles in maximum and minimum gradient directions for 2 m vegetation fraction measures from five representative areas in central Manhattan. In both the forested areas of Central Park and the isolated vegetated courtyards, the characteristic scale of the vegetation is between 5 and 20 m. If the vegetation fraction is spatially homogenous at scales larger than 50 m, pixel registration is not important but in the common case of isolated stands of vegetation with dimensions smaller than the 50 m GIFOV, the response-weighted vegetation fraction is sensitive to small differences in the location of the GIFOV. For this reason it is not generally possible to validate vegetation fraction estimates for individual Landsat pixels unless they occur within larger areas of homogeneous vegetation fraction.

10. Validation

It is evident from figure 7 that there is a strong correspondence between the unmixed Landsat vegetation fraction estimates and those derived from the higher resolution airphoto. Aside from the dropped scan lines, there appears to be good agreement between the Landsat vegetation fraction image and the airphoto-derived estimate, even given the finer sampling of the airphoto estimate. For most applications, regional differences in vegetation fraction will be of greater interest than the fraction of an individual pixel. It may therefore be possible to validate the satellite-derived estimates by comparing the distribution in a well defined area to the distribution derived from the higher resolution airphoto.

If the effect of pixel registration uncertainty is to redistribute vegetation fraction among adjacent Landsat pixels then the overall distribution of vegetation fraction estimates for a region should not generally depend on registration to the extent that an individual pixel estimate does. If the mixing model was consistently over or underestimating the actual vegetation fraction, then the distribution of satellite-derived estimates for a specific area would be consistently different from the distribution of high-resolution estimates. Since individual features with well defined rectangular boundaries are easily recognized in both the satellite and airphoto

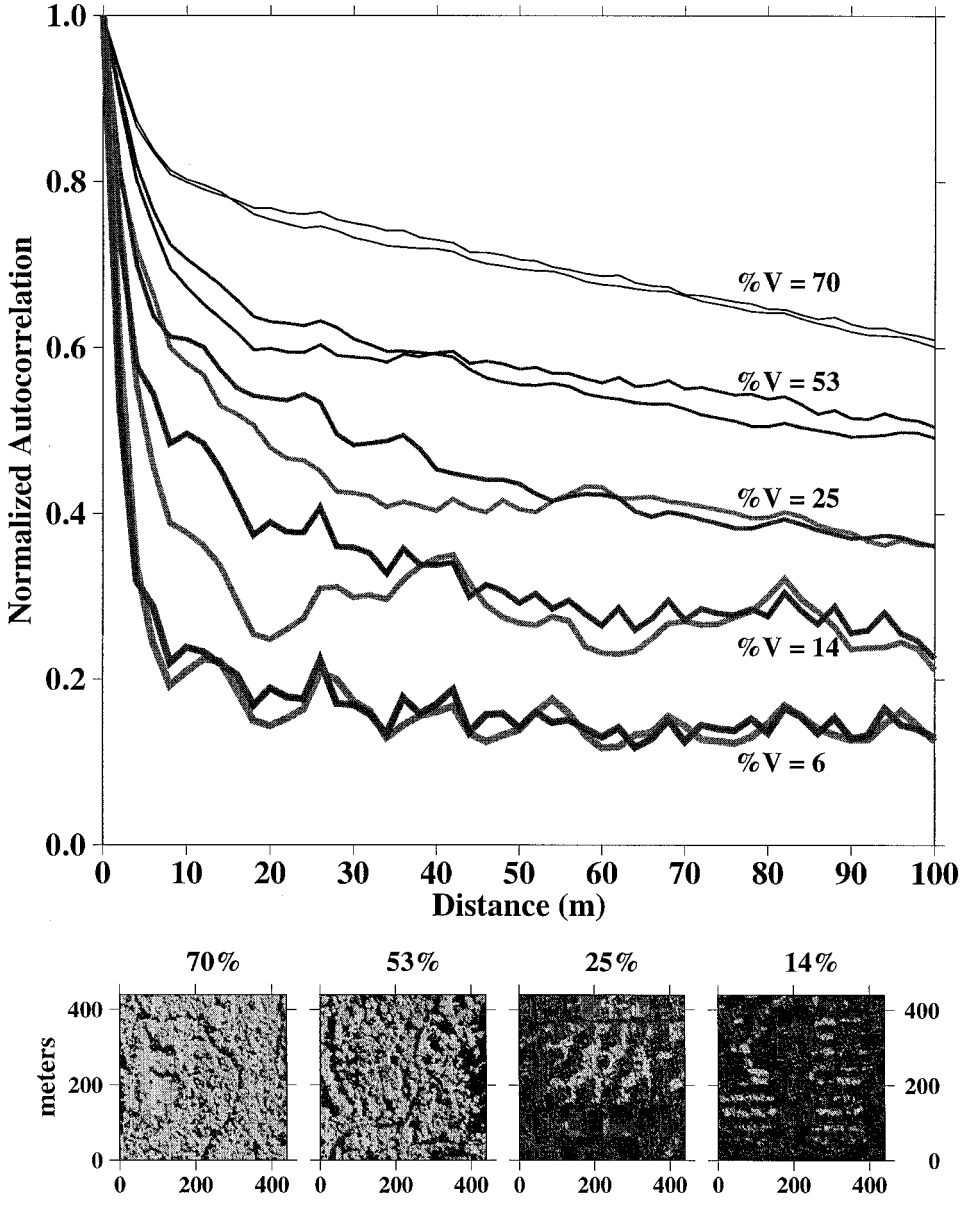


Figure 8. Characteristic scales of urban vegetation. Abrupt breaks in slope of the normalized autocorrelation functions for five representative areas in Manhattan indicate that the characteristic scale of vegetation varies from <10 m to 30 m and is generally smaller than the IFOV of the Landsat TM sensor. The curves represent the maximum gradient (light) and minimum gradient (dark) radial autocorrelation functions for open ($F_V = 0.53$) and closed ($F_V = 0.7$) canopy forest in Central Park, courtyard trees in a housing complex in Harlem ($F_V = 0.25$), intrablock courtyard trees on the Upper West Side ($F_V = 0.14$) and an isolated community garden on the Upper East Side ($F_V = 0.06$). Subpixel registration uncertainty is relatively insignificant for homogeneous vegetated areas but can result in considerable validation error when F_V changes at subpixel scales as with smaller isolated vegetated areas.

estimates in figure 8, it is straightforward to compare distributions for a variety of easily discernible features and areas. Figure 9 compares vegetation fraction distributions for two contrasting regions of central Manhattan as well as for the entire validation region. The distributions from both estimates show good agreement for vegetation fractions greater than 0.2 but significant disagreement for lower fractions. The disagreement for fractions less than 0.2 is a result of a larger modal vegetation fraction for the airphoto distribution than for the satellite distribution. This is at least partially a result of non-vegetated 2 m pixels with small visible green components such as the water bodies which have near-zero vegetation fractions in the satellite estimates but small, non-zero green components in the colour visible airphoto. In this respect, the non-zero uncertainty level of the high-resolution vegetation fraction estimate imposes a limitation on the smallest (<0.2) vegetation fraction that can be validated with the airphoto.

Comparing distributions of measured and unmixed vegetation fractions provides two complementary measures to assess agreement. While similar mean values is a necessary but not sufficient condition for equality of distributions, similar dispersions (e.g. standard deviations) strengthen the assertion that the two distributions are equivalent. Thirty four validation sites of varying size and vegetation distribution were chosen within the central Manhattan area (table 3 and figure 7). The distributions of measured and unmixed vegetation fractions are compared for each area in figure 10. Vegetation fraction distributions are remarkably consistent between the two methods as indicated by nearly equal mean values and standard deviations over nearly the entire range of vegetation fractions. Mean values generally agree to within $F_V < 0.1$ and almost all of the distribution means lie within one standard deviation of perfect agreement. The uniform scatter about the diagonal suggests that there is no systematic bias between the estimates measured from the airphoto and the estimates resulting from inversion of the mixing model. While good agreement would be expected for very high (>0.9) and very low (<0.1) fractions, it is remarkable that low to intermediate fractions agree as closely as they do.

The validation exercise described here should be considered a first step rather than an absolute verification of the spectral mixing model. The flaws and biases inherent in a comparison of a colour visible airphoto with visible/infrared satellite imagery are obvious and should be given serious consideration. The purpose of this validation is to demonstrate that two indirect methods of estimating vegetation fractions in urban environments produce consistent results and that the spectral unmixing approach warrants further investigation. A validation based on simultaneously acquired, high-resolution (2 m) hyperspectral imagery with coincident field measurements would provide more meaningful results.

11. Discussion

The results of this analysis suggest that inversion of a three-component linear spectral mixing model for Landsat TM data can provide reasonable quantitative areal estimates of vegetation fraction over a wide range of abundances in an urban setting. The validation described above is obviously incomplete in that it relies entirely on visible reflectance in the high-resolution image and does not address issues of spatial scaling. The inability of the validation method to quantify the uncertainty of vegetation fraction estimates of less than 0.2 is a significant issue that must be resolved in future studies. Most parks and green spaces in New York City have vegetation fractions larger than 0.2 and are well resolved in the unmixed

Landsat data but accurate estimates of lower vegetation fractions will be important for many applications. In spite of the acknowledged inability of the validation method to constrain these low vegetation abundances, the significant increase in RMS misfit for unmixed vegetation fractions less than 0.2 suggests that this may be a real limitation imposed by the satellite data. In areas with little, or poorly illuminated, vegetation the reflectance is dominated by the other two endmembers. Since the low- and high-albedo endmembers are characterized by a higher dimensional (figure 2) and possibly nonlinear mixing, the three-component mixing model is evidently inappropriate for estimation these endmembers. The influence of atmospheric scattering on both spatial and spectral resolution of optical imagery is known to be significant (Kaufman 1989) so the limits implied by these results should not be assumed to be universal. Atmospheric conditions were far from ideal when this Landsat image was acquired so it may be possible to obtain smaller misfits for low abundances under better conditions or by applying atmospheric corrections. Illumination conditions will also limit the accuracy of reflectance-based estimates of vegetation abundance. Small-scale vegetation that is not illuminated during image acquisition will not be detected by satellites in Sun synchronous orbits yet may amount to a significant fraction of vegetation present in urban areas. Full consideration of spatial scaling issues is beyond the scope of this study. Nonetheless, the surprisingly good agreement between two completely independent estimates over the wide range of vegetation abundances is encouraging and suggests that the method warrants more detailed investigation.

One of the primary advantages of spectral unmixing compared to most commonly used vegetation indices is its relatively straightforward physical interpretation. The apparent first-order linearity of macroscopic spectral mixing suggests a direct correspondence between unmixed vegetation fraction and the area within the GIFOV containing vegetation cover. In contrast, the commonly used NDVI infers the presence of vegetation on the basis of the difference between visible and VNIR reflectance but does not provide areal estimates of the amount of vegetation. In addition to the dependence of NDVI on the spectral bandwidth of the particular sensor, the relationship between NDVI and other measures of vegetation abundance (e.g. Leaf Area Index (LAI)) is notoriously nonlinear (Asrar *et al.* 1984, 1989). This nonlinearity and platform dependence limits the extent to which NDVI can be used for quantitative analyses of vegetation. Figure 11 indicates that the NDVI obtained from Landsat TM is asymptotic above intermediate vegetation fraction in much the same manner that it saturates with increased values of LAI (Hinzman *et al.* 1984). In contrast to the unmixed vegetation fraction, the NDVI does not distinguish between open and closed canopy forest and grass-covered areas in Central Park. The NDVI also overestimates the abundance of interspersed non-park vegetation relative to more densely vegetated areas in parks.

It is not surprising that spectral unmixing provides a better estimate of vegetation

Figure 9. Distribution of measured and unmixed vegetation fractions in central Manhattan. The greatest disparity between the 28.5 m Landsat estimate (bars) and the 2 m airphoto estimate (curves) distributions is seen at low vegetation fractions with generally good agreement for fractions larger than 0.2. The presence of small amounts of non-vegetative visible green reflectance in the airphoto is responsible for the displaced peaks in the distributions around $F_V = 0.1$ and results in an uncertainty of ~ 0.05 in the F_V estimated from the airphoto.

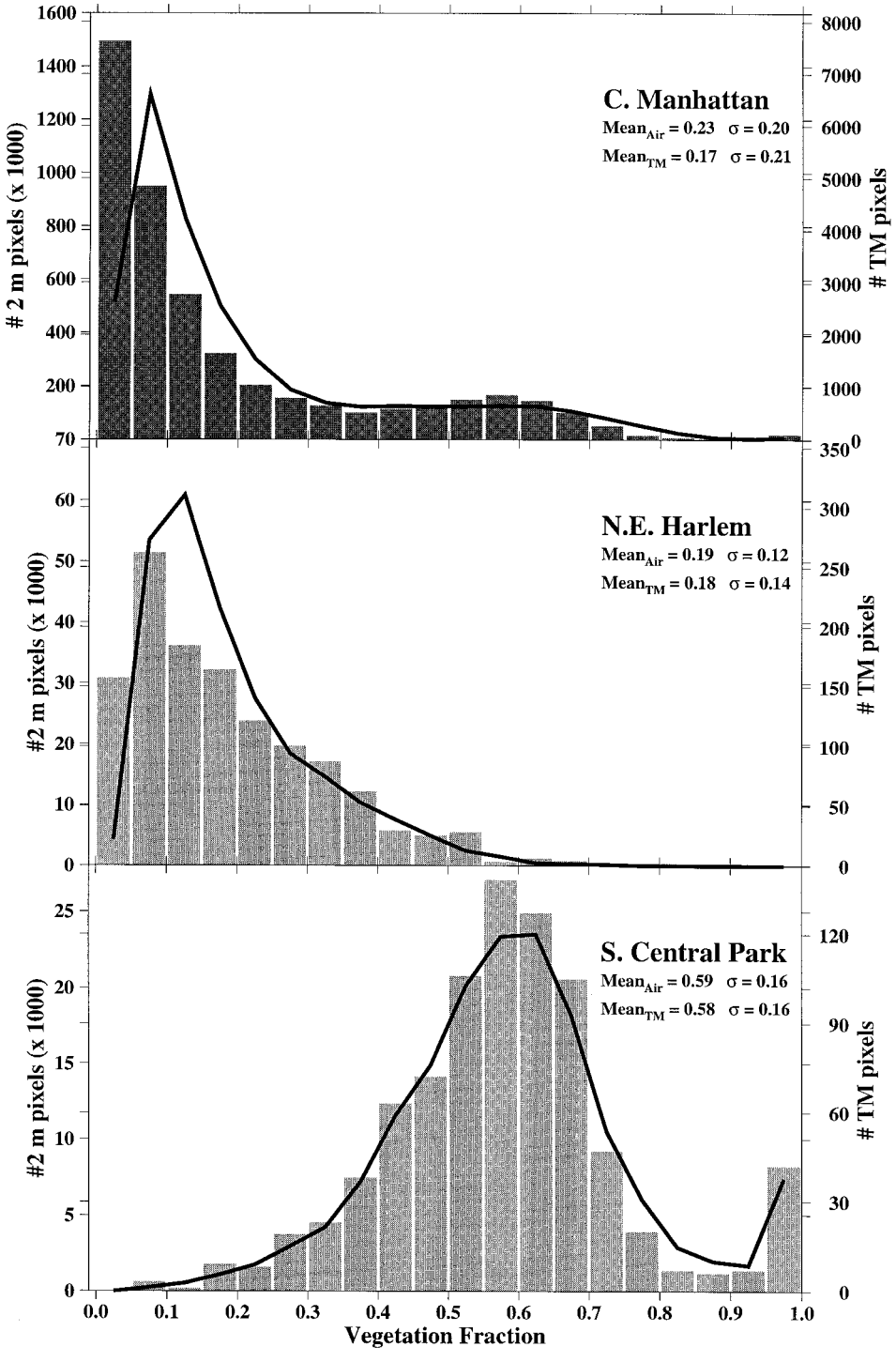

C. Manhattan
Mean$_{Air}$ = 0.23 σ = 0.20
Mean$_{TM}$ = 0.17 σ = 0.21
2 m pixels (x 1000)
TM pixels
N.E. Harlem
Mean$_{Air}$ = 0.19 σ = 0.12
Mean$_{TM}$ = 0.18 σ = 0.14
#2 m pixels (x 1000)
TM pixels
S. Central Park
Mean$_{Air}$ = 0.59 σ = 0.16
Mean$_{TM}$ = 0.58 σ = 0.16
#2 m pixels (x 1000)
TM pixels
Vegetation Fraction

Table 3. Central Manhattan validation sites.

		Measured F_V			Unmixed F_V		
	Site	Mean	σ	N @ 2 m	Mean	σ	N @ 28.5 m
1	Sakura Park	0.48	0.15	3700	0.42	0.16	15
2	Columbia University	0.29	0.14	15 543	0.25	0.18	77
3	St. John's Park	0.34	0.17	15 138	0.26	0.18	70
4	Morningside Park	0.63	0.17	42 705	0.48	0.18	214
5	Courtyard trees	0.45	0.11	641	0.55	0.21	3
6	Courtyard trees	0.35	0.05	377	0.28	0.00	2
7	Community Garden	0.34	0.07	494	0.42	0.06	2
8	Community Garden	0.39	0.09	723	0.23	0.05	6
9	Marcus Garvey Park	0.50	0.15	23 746	0.39	0.16	121
10	Mt. Morris district	0.32	0.10	15 977	0.18	0.12	84
11	M.L.K. Jr. Housing Complex	0.38	0.13	14 659	0.29	0.14	77
12	Ballfield	0.37	0.12	1606	0.44	0.15	6
13	Grover Washington Housing Complex	0.23	0.09	33 048	0.22	0.14	180
14	Ballfield	0.37	0.16	2592	0.36	0.24	15
15	Community garden	0.36	0.08	1404	0.46	0.13	9
16	Cooper Hewitt Museum	0.25	0.11	900	0.34	0.07	4
17	Carl Schurz Park	0.28	0.07	16 131	0.39	0.21	92
18	Amsterdam Housing Complex	0.31	0.16	11 881	0.20	0.15	63
19	American Museum of Natural History	0.26	0.16	21 626	0.30	0.21	110
20	Park West Village	0.42	0.07	1800	0.33	0.07	9
21	North Quarter, Central Park	0.61	0.16	182 320	0.53	0.19	930
22	South Quarter, Central Park	0.56	0.21	182 450	0.51	0.21	928
23	Sheep Meadow, Central Park	0.99	0.02	5525	0.99	0.03	25
24	The Ballground, Central Park	0.77	0.14	3078	0.70	0.09	15
25	Courtyard trees	0.11	0.05	15 872	0.10	0.08	81
26	Courtyard trees	0.25	0.05	646	0.38	0.01	2
27	Courtyard trees	0.24	0.04	496	0.24	0.01	2
28	Courtyard trees	0.40	0.03	195	0.31	0.00	1
29	Courtyard trees	0.62	0.04	238	0.56	0.00	1
30	Community garden	0.44	0.06	255	0.49	0.00	1
31	Randall's Island Field	0.97	0.04	1419	0.93	0.03	6
32	Great Hill, Central Park	0.69	0.14	35 424	0.58	0.16	180
33	Cedar Hill, Central Park	0.72	0.09	1242	0.82	0.08	6
34	The Ramble, Central Park	0.62	0.07	3975	0.67	0.12	20

abundance than does the NDVI. Estimation of vegetation fraction from linear spectral unmixing depends not only on the amplitude of the visible/VNIR 'red edge' but on the overall shape of the spectral reflectance profile over a larger portion of the spectrum. Fraction estimates obtained by inverting Landsat TM reflectances provide a more quantitative, physically based measure of vegetation abundance that may be more easily used to constrain quantities like evapotranspiration and biomass. For applications that require biophysical measures, such as LAI or biomass, additional studies must be carried out to calibrate the relationship between the biophysical measure and the vegetation fraction in the urban setting.

Vegetation estimates that make use of the full spectral reflectance vector should show better agreement with the unmixed vegetation fraction. This is indeed the case as the Landsat TM Tasseled Cap Greenness (Crist and Ciccone 1984) shows a very

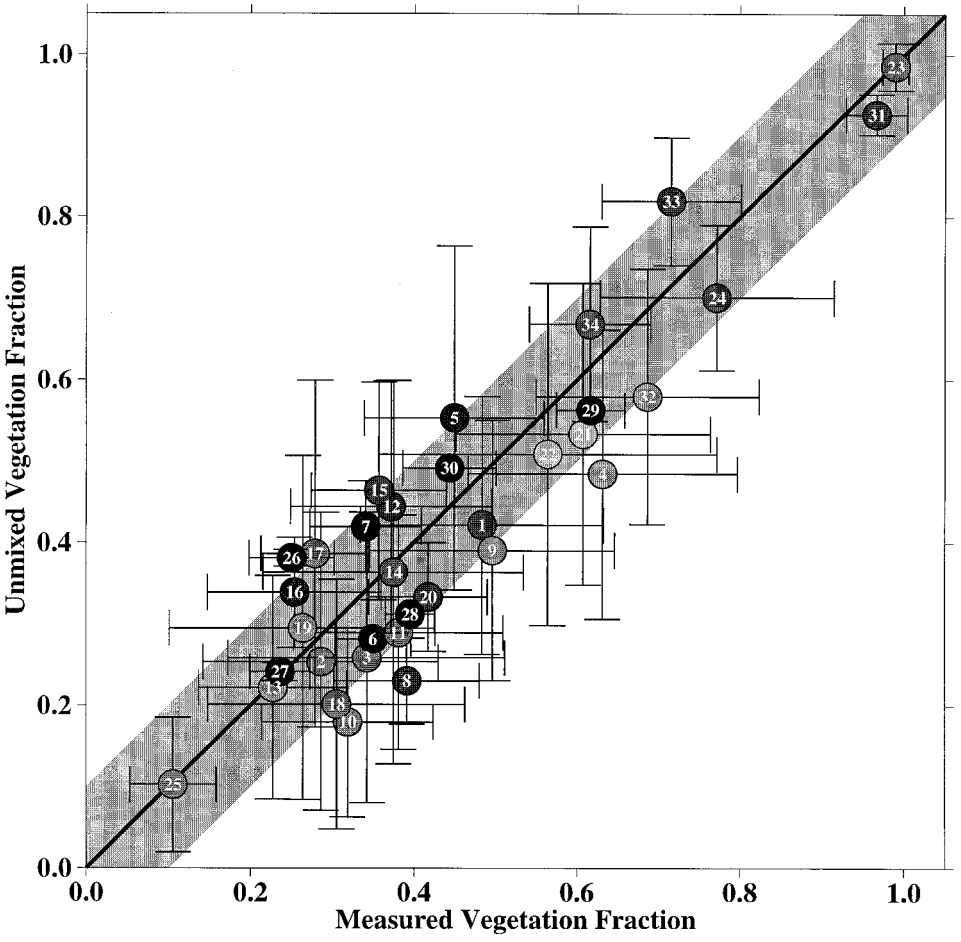


Figure 10. Validation of the linear mixture model. Mean unmixed and measured F_V for 34 validation sites in central Manhattan generally agree to within ± 0.1 (grey diagonal) while the standard deviations (bars) of the distributions of measured and unmixed F_V are generally greater than 0.1. The sizes of the validation sites are indicated by the shading of the symbols ranging from less than 10 (dark) to ~ 1000 (light) TM pixels. The numbers correspond to locations of validation sites, see table 3 and figure 7. The linear correlation coefficient of the measured and unmixed fractions is 0.999.

strong (0.98) correlation with the unmixed vegetation fraction for the New York City subscene discussed here. In spite of the good agreement, the unmixed vegetation fraction would generally be considered preferable to Tasseled Cap greenness because it provides a quantitative measure of areal vegetation abundance within mixed pixels and incorporates scene specific components of the observed reflectance.

In conclusion, the linear spectral mixing model provides a simple, physically based measure of vegetation abundance and distribution in urban areas. Both constrained and unconstrained inversions of three independently derived ensembles of endmembers produce nearly identical estimates of F_V, suggesting that the inversion is stable and well posed. Quantitative validation with higher spatial resolution (2 m) vegetation estimates derived from aerial photography suggests agreement to within 0.1 for vegetation fractions greater than 0.2. The accuracy of estimates for smaller

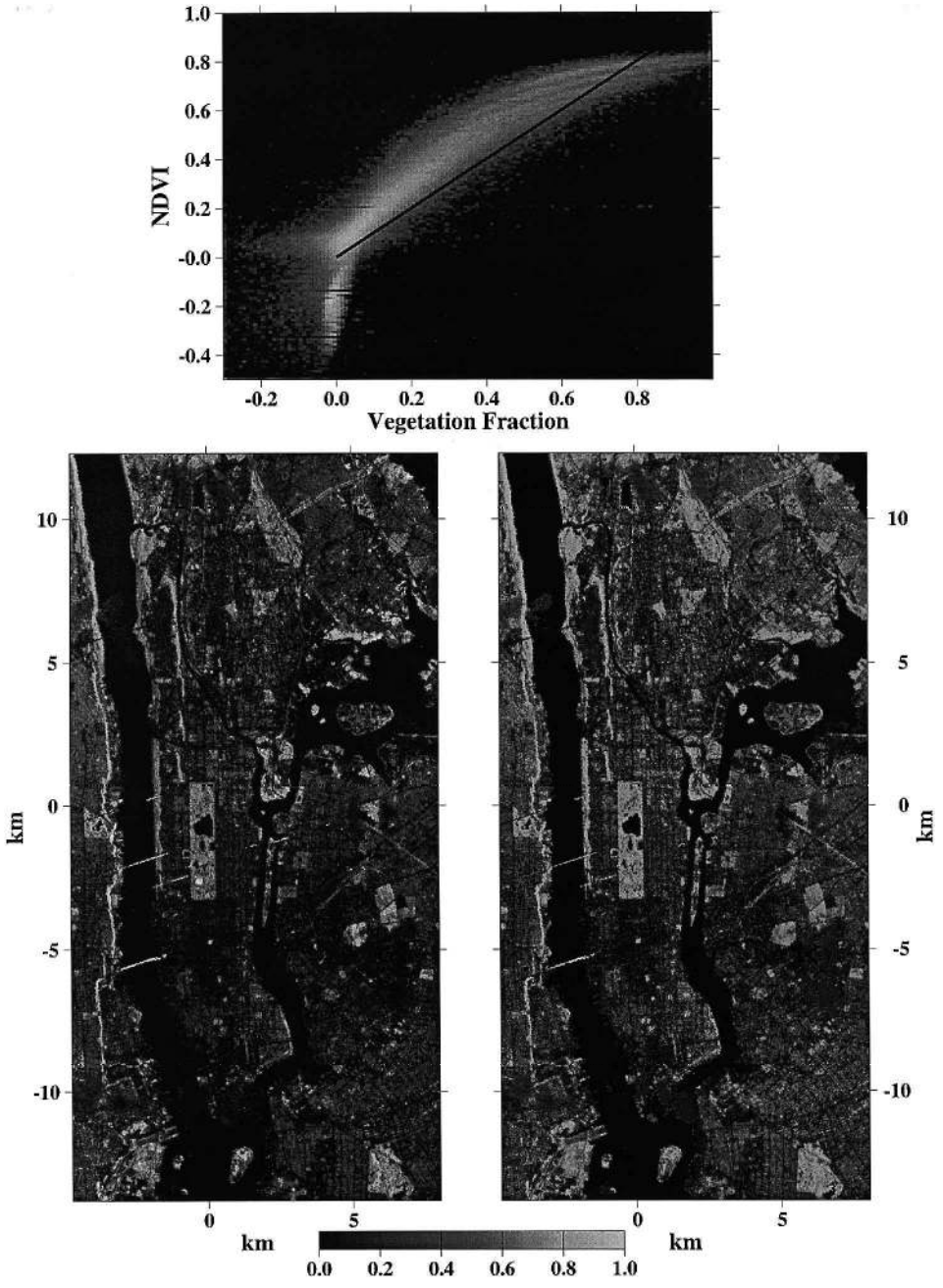


Figure 11. Comparison of unmixed vegetation fraction and NDVI for Manhattan and surrounding areas. The image on the lower left indicates that the unmixed F_V discriminates between grass and forest with varying degrees of canopy closure (e.g. Central Park) while revealing significant intra-urban variations in non-park vegetation abundance. The NDVI image (lower right) does not discriminate canopy closure differences or the presence of grass. The density-shaded scattergram (top) shows some correlation for low values of F_V and NDVI but a pronounced saturation of NDVI for F_V greater than $\sim$0.4. As a result, NDVI values greater than $\sim$0.4 carry an uncertainty of at least 0.4 in vegetation abundance.

vegetation fraction and the limitations imposed by atmospheric scattering and illumination conditions remain to be determined. A validation based on simultaneously acquired, high-resolution (2 m) hyperspectral imagery with coincident field measurements would provide more meaningful results. Studies in other urban areas are necessary to assess the general applicability of the three-component mixing model. Areas with a greater diversity of landcover types may require more than three endmembers. In some cases, a larger number of landcover types may cause the unmixing problem to be underdetermined with Landsat TM data. In spite of these limitations, the results of this study suggest that spectral unmixing may provide simple, physically meaningful estimates of urban vegetation abundance and that the method warrants further investigation.

Acknowledgments

This research was supported by the University Consortium for Atmospheric Research (UCAR) and the Columbia Earth Institute. The author thanks Zhikang Chen and two anonymous reviewers for valuable comments and suggestions. Landsat data from the NASA Landsat Data Collection were provided by the USGS EROS Data Center, Sioux Falls, SD. Digital aerial photography data were provided by Pacific Aerial Surveys, Hammon, Jensen, Wallen and Associates Inc. Much of the analysis presented here was performed using ENVI (v3.0) image processing software and Matlab numerical analysis software. Figures were prepared using GMT software. Computing facilities were funded by the NASA Centers of Excellence in Remote Sensing program.

References

Adams, J. B., Sabol, D. E., Kapos, V., Filho, R. A., Roberts, D. A., Smith, M. O., and Gillespie, A. R., 1995, Classification of multispectral images based on fractions of endmembers: application to landcover change in the Brazilian Amazon. *Remote Sensing of Environment*, **52**, 137–154.

Adams, J. B., Smith, M. O., and Johnson, P. E., 1986, Spectral mixture modelling: a new analysis of rock and soil types at the Viking Lander 1 site. *Journal of Geophysical Research*, **91**, 8098–8122.

Asner, G. P., 1998, Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment*, **64**, 234–253.

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

Asrar, G., Myneni, R. B., and Kanematsu, E. T., 1989, Estimation of plant canopy attributes from spectral reflectance measurements. In *Theory and Applications of Optical Remote Sensing*, edited by G. Asrar (New York: Wiley).

Berry, B. L., 1990, Urbanization. In *The Earth as Transformed by Human Action*, edited by B. L. Turner II *et al.* (Cambridge: Cambridge University Press), pp. 103–119.

Boardman, J. W., 1989, Inversion of imaging spectrometry data using singular value decomposition. *Proceedings of IGARSS'89, 12th Canadian Symposium on Remote Sensing, Calgary, Canada*, pp. 2069–2072.

Boardman, J. W., 1993, Automating spectral unmixing of AVIRIS data using convex geometry concepts. *Proceedings of the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Airborne Geoscience Workshop, Jet Propulsion Laboratory, Pasadena, CA, 25–29 October 1993* (Pasadena, CA: JPL), pp. 11–14.

Boardman, J. W., and Kruse, F. A., 1994, Automated spectral analysis: a geologic example using AVIRIS data, north Grapevine mountains, Nevada. *Proceedings of the Tenth Thematic Conference on Geologic Remote Sensing, Ann Arbor, MI* (Ann Arbor, MI: ERIM), pp. I407–I418.

Borel, C. C., and Gerstl, S. A. W., 1994, Nonlinear spectral mixing models for vegetative and soil surfaces. *Remote Sensing of Environment*, **47**, 403–416.

Carlson, T. N., Gillies, R. R., and Perry, E. M., 1994, A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover. *Remote Sensing Reviews*, **9**, 161-173.

Clark, R. N., and Lucey, P. G., 1984, Spectral properties of ice-particulate mixtures and implications for remote sensing 1. Intimate mixtures. *Journal of Geophysical Research*, **89**, 6341–6348.

Clark, R. N., and Roush, T. L., 1984, Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research*, **89**, 6329–6340.

Clark, R. N., Swayze, G. A., Gallagher, A. J., King, T. V. V., and Calvin, W. M., 1993, Digital Spectral Library: Version 1: 0.2 to 3.0 microns. U. S. Geological Survey, 93–592.

Crist, E. P., and Cicone, R. C., 1984, A physically-based transformation of thematic mapper data—The TM tasseled cap. *IEEE Transactions on Geoscience and Remote Sensing*, **GE-22**, 256–263.

Elvidge, C. D., Chen, Z., and Groeneveld, D. P., 1993, Detection of trace quantities of green vegetation in 1990 AVIRIS data. *Remote Sensing of Environment*, **44**, 271–279.

Forster, B., 1983, Some urban measurements from Landsat data. *Photogrammetric Engineering and Remote Sensing*, **49**, 1293–1707.

Forster, B. C., 1985, An examination of some problems and solutions in monitoring urban areas from satellite platforms. *International Journal of Remote Sensing*, **6**, 139–151.

Gallo, K. P., McNab, A. L., Karl, T. R., Brown, J. F., Hood, J. J., and Tarpley, J. D., 1993, The use of a vegetation index for assessment of the urban heat island effect. *International Journal of Remote Sensing*, **14**, 2223–2230.

Gillespie, A. R., Smith, M. D., Adams, J. B., Willis, S. C., Fischer, A. F., and Sabol, D. E., 1990, Interpretation of residual images: spectral mixture analysis of AVIRIS. *Proceedings of the 2nd AVIRIS Workshop, 4–5 June 1990* (Pasadena, CA: JPL), pp. 243–270.

Gillies, R. R., Carlson, T. N., Cui, J., Kustas, W. P., and Humes, K. S., 1997, A verification of the 'triangle' method for obtaining surface soil water content and energy fluxes from remote measurements of the Normalized Difference Vegetation Index (NDVI) and surface radiant temperature. *International Journal of Remote Sensing*, **18**, 3145–3166.

Goetz, A. F. H., Vane, G., Solomon, J. E., and Rock, B. N., 1985, Imaging spectrometry for earth remote sensing. *Science*, **228**, 1147–1153.

Goward, S. N., Cruickshanks, G. D., and Hope, A. S., 1985, Observed relation between thermal emission and reflected spectral radiance of a complex vegetated landscape. *Remote Sensing of Environment*, **18**, 137–146.

Green, A. A., Berman, M., Switzer, P., and Craig, M. D., 1988, A transformation for ordering mutispecral data in terms of image quality with implications for noise removal. *IEEE Transactions on Geoscience and Remote Sensing*, **26**, 65–74.

Hinzman, L. D., Bauer, M. E., and Daughtry, C. S., 1984, Growth and reflectance characteristics of winter wheat canopies. *Technical Report, NASA/JSC*, 111484.

Horowitz, H. M., Nalepka, R. F., Hyde, P. D., and Morganstern, J. P., 1971, Estimating the proportions of objects within a single resolution element of a multispectral scanner. University of Michigan, NASA contract NAS-9-9784.

Jensen, J. R., 1996, *Introductory Digital Image Processing*, 2nd edition (Upper Saddle River, NJ: Prentice Hall).

Kaufman, Y. J., 1989, The atmospheric effect on remote sensing and its corrections. In *Theory and Applications of Optical Remote Sensing*, edited by G. Asrar (New York: Wiley), pp. 336–428.

Kauth, R. J., and Thomas, G. S., 1976, The Tasseled Cap—a graphic description of the spectral–temporal development of agricultural crops as seen by Landsat. *Proceedings of the Symposium on Machine Porcessing of Remotely Sensed Data, Purdue University, West Lafayette, Indiana*, pp. 4B41–4B51.

Kruse, F. A., 1988, Use of airborne imaging spectrometer data to map minerals associated with hydrotermally altered rocks in the northern Grapevine Mountains, Nevada and California. *Remote Sensing of Environment*, **24**, 31–51.

Landsberg, H. E., 1981, *The Urban Climate* (New York: Academic).
Lee, J. B., Woodyatt, A. S., and Berman, M., 1990, Enhancement of high specral resolution remote sensing data by a noise-adjusted principal components tranform. *IEEE Transactions on Geoscience and Remote Sensing*, **28**, 295–304.
Markham, B. L., 1985, The Landsat sensors' spatial responses. *IEEE Transactions on Geoscience and Remote Sensing*, **23**, 864–875.
Markham, B. L., and Barker, J. L., 1987, Thematic Mapper bandpass solar exoatmospheric irradiances. *International Journal of Remote Sensing*, **8**, 517–523.
Menke, W., 1989, *Geophysical Data Analysis: Discrete Inverse Theory* (New York: Academic).
National Geographic Society, 1998, *National Geographic Satellite Atlas of the World* (Washington, DC: National Geographic Society).
Nemani, R. R., and Running, S. W., 1989, Estimation of regional surface resistance to evapotranspiration from NDVI and thermal-IR AVHRR data. *Journal of Applied Meteorology*, **28**, 276–284.
Nichol, J. E., 1996, High-resolution surface temperature patterns related to urban morphology in a tropical city; a satellite-based study. *Journal of Applied Meteorology*, **28**, 276–284.
Oke, T. R., 1982, The energetic basis of the urban heat island. *Quarterly Journal of the Royal Meteorological Society*, **108**, 1–24.
Ormsby, J. P., 1992, Evaluation of natural and man-made features using Landsat TM data. *International Journal of Remote Sensing*, **13**, 303–318.
Owen, T. W., Carlson, T. N., and Gilllies, R. R., 1998, An assessment of satellite remotely-sensed landcover parameters in quantitatively describing the climatic effect of urbanization. *International Journal of Remote Sensing*, **19**, 1663–1681.
Parker, R. L., 1994, *Geophysical Inverse Theory* (Princeton, NJ: Princeton University Press).
Pech, R. P., Davies, A. W., Lamacraft, R. R., and Graetz, R. D., 1986, Calibaration of Landsat data for sparsely vegetated semi-arid rangelands. *International Journal of Remote Sensing*, **7**, 1729–1750.
Preisendorfer, R. W., 1988, In *Principal Component Analysis in Meteorology and Oceanography*, edited by C. D. Mobley (Amsterdam: Elsevier).
Price, J. C., 1990, Using spatial context in satellite data to infer regional scale evapotranspiration. *IEEE Transactions on Geoscience and Remote Sensing*, **28**, 940–948.
Richards, J. A., 1993, *Remote Sensing Digital Analysis: An Introduction* (New York: Springer-Verlag).
Roberts, D. A., Smith, M. O., and Adams, J. B., 1993, Green vegetation, nonphotosynthetic vegetation and soils in AVIRIS data. *Remote Sensing of Environment*, **44**, 255–269.
Roth, M., Oke, T. R., and Emery, W. J., 1989, Satellite-derived urban heat islands from three coastal cities and the utilization of such data in urban climatology. *International Journal of Remote Sensing*, **10**, 1699–1720.
Settle, J. J., and Drake, N. A., 1993, Linear mixing and the estimation of ground cover proportions. *International Journal of Remote Sensing*, **14**, 1159–1177.
Singer, R. B., and McCord, T. B., 1979, Mars: large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance. *Proceedings of 10th Lunar and Planetary Science Conference, American Geophysical Union, Washington, DC*, pp. 1835–1848.
Smith, M. O., Johnson, P. E., and Adams, J. B., 1985, Quantitative determination of mineral types and abundances from reflectance spectra using principal component analysis. *Journal of Geophysical Research*, **90**, 792–804.
Smith, M. O., Ustin, S. L., Adams, J. B., and Gillespie, A. R., 1990, Vegetation in deserts: I. A regional measure of abundance from multispectral images, *Remote Sensing of Environment*, **31**, 1–26.
Strang, G., 1986, *Introduction to Applied Mathematics* (Cambridge, MA: Wellesley).
United Nations, 1980, Urban, Rural and City Population, 1950-2000. Department of International Economic and Social Affairs, Population Division, ESA/P/PW.66.
United Nations, 1997, Prospects for Urbanization. ST/ESA/SER.A/166, Sales No. E.97.XIII.3.
Wagrowski, D. M., and Hites, R. A., 1997, Polycyclic aromatic hydrocarbon accumulation in urban, suburban and rurual vegetation. *Environmental Science and Technology*, **31**, 279–282.
Welch, R., 1982, Spatial resolution requirements for urban studies. *International Journal of Remote Sensing*, **3**, 139–146.

Wessman, C. A., Bateson, C. A., Curtiss, B., and Benning, T. L., 1994, A comparison of spectral mixture analysis and NDVI for ascertaining ecological variables. *Proceedings of the Seventh Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) Workshop, Jet Propulsion Laboratory, Pasadena, CA*.

Wilson, A. K., 1988, The effective resolution element of Landsat Thematic Mapper. *International Journal of Remote Sensing*, **9**, 1303–1314.

Woodcock, C. E., and Strahler, A. H., 1987, The factor of scale in remote sensing. *Remote Sensing of Environment*, **21**, 311–322.