

An Improved Dark-Object Subtraction Technique for Atmospheric Scattering Correction of Multispectral Data*

PAT S. CHAVEZ, JR.

U.S. Geological Survey, 2255 N. Gemini Drive, Flagstaff, Arizona 86001

Digital analysis of remotely sensed data has become an important component of many earth-science studies. These data are often processed through a set of preprocessing or "clean-up" routines that includes a correction for atmospheric scattering, often called haze. Various methods to correct or remove the additive haze component have been developed, including the widely used dark-object subtraction technique. A problem with most of these methods is that the haze values for each spectral band are selected independently. This can create problems because atmospheric scattering is highly wavelength-dependent in the visible part of the electromagnetic spectrum and the scattering values are correlated with each other. Therefore, multispectral data such as from the Landsat Thematic Mapper and Multispectral Scanner must be corrected with haze values that are spectral band dependent. An improved dark-object subtraction technique is demonstrated that allows the user to select a *relative* atmospheric scattering model to predict the haze values for all the spectral bands from a selected starting band haze value. The improved method normalizes the predicted haze values for the different gain and offset parameters used by the imaging system. Examples of haze value differences between the old and improved methods for Thematic Mapper Bands 1, 2, 3, 4, 5, and 7 are 40.0, 13.0, 12.0, 8.0, 5.0, and 2.0 vs. 40.0, 13.2, 8.9, 4.9, 16.7, and 3.3, respectively, using a relative scattering model of a clear atmosphere. In one Landsat multispectral scanner image the haze value differences for Bands 4, 5, 6, and 7 were 30.0, 50.0, 50.0, and 40.0 for the old method vs. 30.0, 34.4, 43.6, and 6.4 for the new method using a relative scattering model of a hazy atmosphere.

Introduction

Digital analysis of remotely sensed data has become an important component of a wide variety of earth science studies. However, regardless of the analysis to be done, most remotely sensed data are processed through a set of preprocessing or clean-up routines that remove or correct unwanted artifacts (both geometric and radiometric). A routine that corrects for atmospheric scattering effects is often included in the radiometric preprocessing.

The atmosphere influences the amount of electromagnetic energy that is sensed by the detectors of an imaging system, and these effects are wavelength depen-

dent (Curcio, 1961; Turner et al., 1971; Sabins, 1978; Slater et al., 1983). This is particularly true for imaging systems such as the Landsat Multispectral Scanner (MSS) and Thematic Mapper (TM) that record data in the visible and near-infrared parts of the spectrum. The atmosphere affects images by scattering, absorbing, and refracting light. The most dominant of these is usually light scattering (Siegel et al., 1980; Slater et al., 1983). Various methods to remove the effects of scattering, which is an additive component, from these data have been developed, including a dark-object subtraction technique. Atmospheric absorption has a multiplicative affect on the data and methods to correct for this particular component are usually not part of a dark-object subtraction technique and are not considered in this paper. The method

* Publication authorized by the Director, U.S. Geological Survey on 4 April 1986.

presented here is strictly an improvement to existing dark-object subtraction techniques that derive the correction DN (digital number) values solely from the digital data with no outside information.

The technique used here allows the user to select a *relative* atmospheric scattering model and predict the haze values, with gain and offset normalization, for the bands being used from a selected starting band haze value. This method should allow better results to be generated from analyses of remotely sensed data for studies where atmospheric scattering corrections are critical.

Review of Haze-Correction Techniques

The effects of atmospheric scattering on remotely sensed multispectral digital image data have been extensively studied and documented during the past 15 years. For example, the amount of scattering in the data can affect results generated from multispectral ratioing (Vincent, 1973; Rowan et al., 1974; Holben and Justice, 1981), estimation of forest leaf area index (Spanner et al., 1984), temporal analyses (Chavez et al., 1977; Otterman and Robinove, 1981), and signature extension (Carr et al., 1983).

Several different atmospheric scattering or "haze" removal techniques have been developed for use with digital remotely sensed data (e.g., Landsat MSS and TM). Most of these techniques can be grouped into a simple dark-object subtraction method (Vincent, 1972; Chavez, 1975; Rowan et al., 1974) or more sophisticated methods that use various atmospheric transmission models, *in situ* field data, or require specific targets to be present in the image (Ahern et al., 1977; Otterman and Robinove, 1981). A major

disadvantage with the more sophisticated techniques is that they require information other than the digital image data [e.g., path radiance and (or) atmospheric transmission at several locations within the image area collected during the satellite's overflight].

Ideally, a method that uses *in situ* or ground-truth information is the most accurate in terms of correcting for atmospheric haze effects. However, most users must work with remotely sensed data that have already been collected. It is under this condition that the simple dark-object subtraction technique can be used because it requires only information contained in the digital image data. This type of correction involves subtracting a constant DN value from the entire digital image. This assumes a constant haze value throughout the entire image, which is often not the case. However, it does accomplish a first-order correction, which is better than no correction at all. A different constant is used for each spectral band, with a different set of constants used from image to image.

Most dark-object subtraction techniques assume that there is a high probability that there are at least a few pixels within an image which should be black (0% reflectance). This assumption is made because in a single band there are a very large number of pixels. For example, for Landsat MSS and TM single-band images there are over 7 million and 45 million pixels, respectively. Thus, there are usually some shadows due to topography or clouds in the image where the pixels should be completely dark (Fig. 1). Ideally, the imaging system should not detect any radiance at these shadow locations, and a DN value of zero should be assigned to them. However, because of

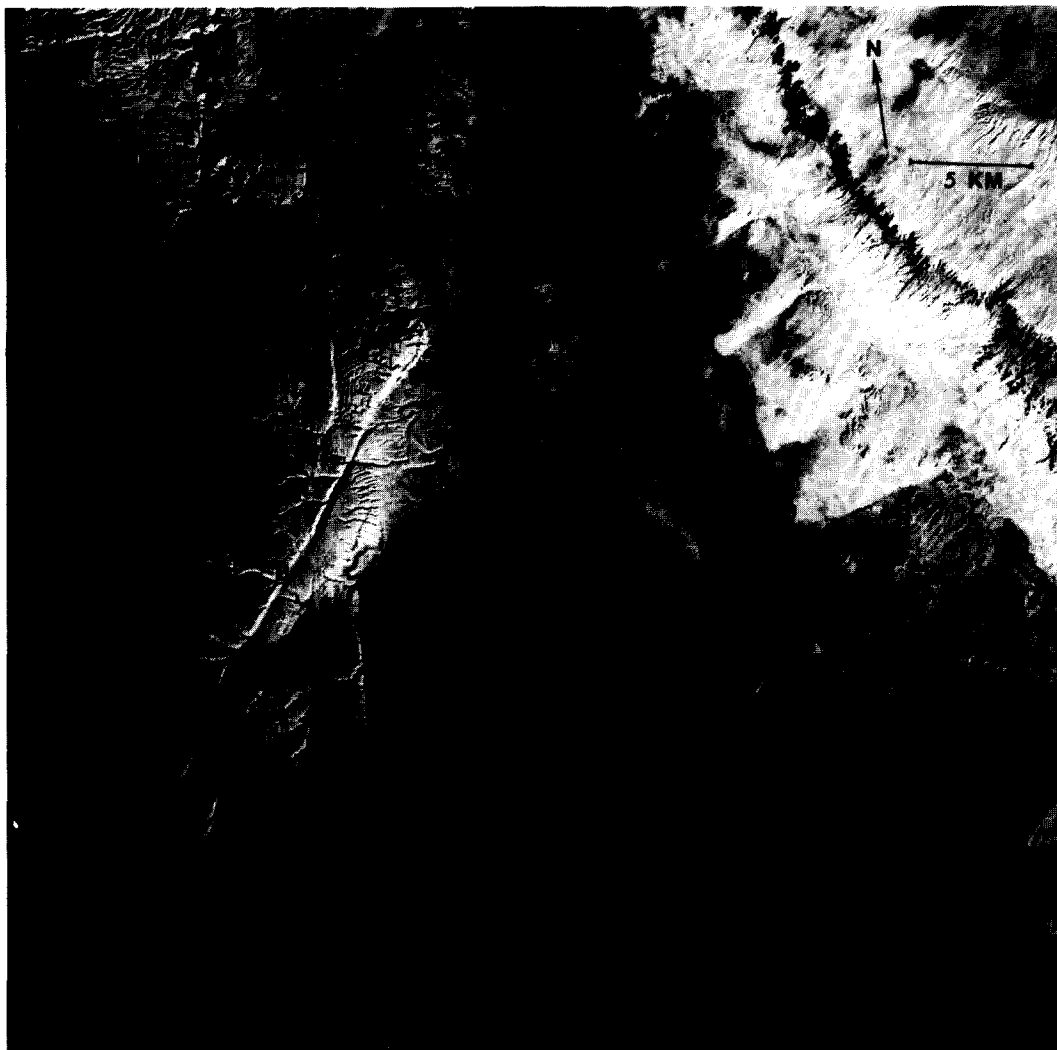


FIGURE 1. Contrasted-stretched Landsat-1 Thematic Mapper (TM) Band 5 image showing a volcanic and sedimentary terrane in northern Arizona. The image was recorded on 21 November 1982. Annotated features include: A) SP lava flow; B) Mesa Butte graben; C) deep shadowed area in the Little Colorado River gorge.

atmospheric scattering effects, these shadowed areas will not be completely dark. Secondary scattering into shadowed areas, such as skylight, is not considered here.

Because of atmospheric scattering, the imaging system records a nonzero DN value at these supposedly dark-shadowed pixel locations. This represents the DN

value that must be subtracted from the particular spectral band to remove the first-order scattering component. Several methods can be used to extract this DN value from the digital data.

A method referred to as the histogram method allows the user to select the haze DN value directly from the DN frequency histogram of a digital image. Histograms

(a)	DN	FREQ.
17	1.	
39	2.	
40	18.	
41	137.	
42	508.	*
43	1094.	*
44	2571.	***
45	5091.	*****
46	7190.	*****
47	9182.	*****
48	11020.	*****
49	13604.	*****
50	15025.	*****
51	15896.	*****
52	18069.	*****
53	23774.	*****
54	29724.	*****
55	31606.	*****
56	35975.	*****
57	43992.	*****
58	51788.	*****
59	55020.	*****
60	55398.	*****
61	58245.	*****
62	70195.	*****
63	81114.	*****
64	78485.	*****
65	69506.	*****
66	74531.	*****
67	65763.	*****
68	56782.	*****
69	57912.	*****
70	57813.	*****
71	57248.	*****
72	51812.	*****
73	45164.	*****
74	43146.	*****
75	41066.	*****
76	33641.	*****
77	26430.	*****
78	25079.	*****

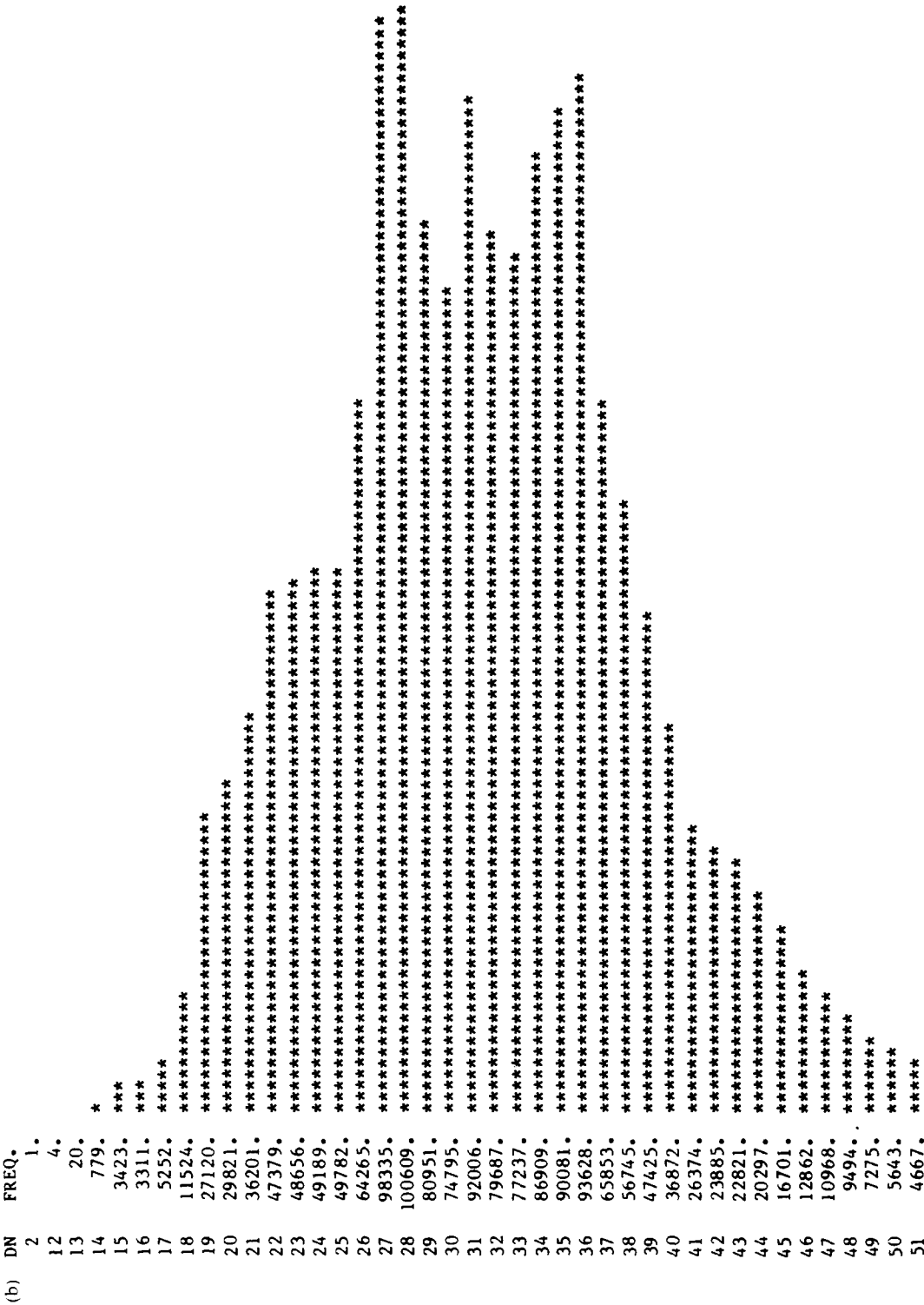


FIGURE 2(a) and 2(b). Lower portion of the DN histograms for Landsat TM Bands 1 (top) and 2 (bottom) of the image shown in Fig. 1. The DN values needed by most dark object subtraction techniques to correct for atmospheric scattering can be selected from the histograms. The DN values for TM Bands 1 and 2 would be 40 and 13, respectively.

(c)

DN	FREQ.
11	1.
12	69.
13	734.
14	3412.
15	2499.
16	2398.
17	4055.
18	9362.
19	13562.
20	14966.
21	19181.
22	26321.
23	25329.
24	24505.
25	22747.
26	29138.
27	30277.
28	31081.
29	31781.
30	43290.
31	41194.
32	38491.
33	39365.
34	44643.
35	48396.
36	43437.
37	41910.
38	47026.
39	47332.
40	41297.
41	40124.
42	40677.
43	55241.
44	54299.
45	56747.
46	52222.
47	61715.
48	46678.
49	42493.
50	37259.
51	39290.

(d)	DN	FREQ.
	0	1.
	8	5.
	9	226.
	10	1421.
	11	3286.
	12	2882.
	13	2757.
	14	3211.
	15	2925.
	16	2961.
	17	4251.
	18	7153.
	19	12182.
	20	12932.
	21	17105.
	22	20066.
	23	18047.
	24	18051.
	25	22064.
	26	36759.
	27	46353.
	28	42446.
	29	40541.
	30	55068.
	31	46846.
	32	42625.
	33	51539.
	34	58712.
	35	58077.
	36	45158.
	37	49937.
	38	55601.
	39	52932.
	40	46201.
	41	48865.
	42	58284.
	43	55517.
	44	58316.
	45	50377.
	46	49791.
	47	48829.

FIGURE 2(c) and 2(d). Lower portions of the DN histograms for Landsat TM bands 3 (top) and 4 (bottom) of the image shown in Fig. 1. The haze correction DN values for TM Bands 3 and 4 would be 12 and 8, respectively.

for given spectral bands, in particular for bands in the visible spectrum, will be offset towards higher DN values by some amount due to scattering. On these histograms there is usually a very sharp increase in the number of pixels at some nonzero DN or grey-level X (Fig. 2). This DN value is assumed to be the amount of haze in that particular band. If this method is used, a histogram must be generated using the entire image, or at least a large portion of the original image. If histograms of very localized areas are used, a *real* minimum or dark-object may not be found, and the DN value selected will overcorrect for atmospheric haze. Using this method on the histograms shown in Figs. 2(a), 2(b), 2(c), and 2(d), which represent Landsat TM Bands 1, 2, 3, and 4 for the area shown in Fig. 1, the DN values selected for the dark-object subtraction haze correction technique are 40, 13, 12, and 8 for TM Bands 1, 2, 3, and 4, respectively.

Improved Method

If digital multispectral image data are haze-corrected by using the standard dark-object subtraction technique described above, problems may be encountered in the analysis stage because the DN values selected for haze removal may not conform to a realistic *relative* atmospheric scattering model. This lack of conformity may cause the data to be overcorrected in some or all of the spectral bands, and the spectral relationship between the bands will not be correct.

In the method proposed here the user selects a starting band dark-object subtraction haze value using the histogram of one of the spectral bands, which is a reasonable way to select the starting

dark-object haze value (Ahern et al., 1982). The user then selects a relative scattering model that he/she feels best represents the atmospheric conditions at the time of data collection. The amplitude of the starting haze can be used as a guide to identify the type of atmospheric conditions that existed during data collection (i.e., very clear, clear, moderate, hazy, or very hazy). The selected relative scattering model is then used to predict the haze values for the other spectral bands from the starting haze value.

Two well-known relative scattering models are the Rayleigh and Mie models (Slater et al., 1983). The Rayleigh model states that relative scattering is inversely proportional to the fourth power of the wavelength (λ^{-4}), which means that shorter wavelengths of the spectrum are scattered much more than the longer wavelengths. This type of scattering is predominantly caused by gas molecules that are much smaller than the wavelengths of light. The Mie model states that scattering is inversely proportional to wavelength, and, in general for a moderate atmosphere, the factor is λ^{-1} . However, this relationship can vary from λ^0 to λ^{-4} , with λ^0 representing complete scattering (e.g., complete cloud cover). Mie scattering is caused by particles that are approximately the same size as the wavelengths, such as smoke and dust particles.

It is suggested that the relative scattering that usually occurs in a *real* atmosphere that is clear seems to follow more of a λ^{-2} to $\lambda^{-0.7}$ relationship and not a Rayleigh or Mie (Curcio, 1961; Slater et al., 1983). Using this information, the relative scattering that occurs in a hazy atmosphere can be approximated as $\lambda^{-0.7}$ to $\lambda^{-0.5}$, if similar power law relation-

ships are used. If more scattering than this occurs, the quality of the data are probably below that desired by the user and a different data set should be used.

The critical aspect of the method proposed in this paper is that haze-correction DN values used by dark-object subtraction techniques be computed using a relative scattering model to ensure that the haze values do represent, or better approximate, true atmospheric scattering possibilities. Using the information supplied by Curcio (1961), and Slater et al. (1983), and extrapolating to very clear and very hazy atmospheres, one possible set of relative scattering models are:

Atmospheric Conditions	Relative Scattering Model
Very clear	λ^{-4}
Clear	λ^{-2}
Moderate	λ^{-1}
Hazy	$\lambda^{-0.7}$
Very hazy	$\lambda^{-0.5}$

From work done by the author using TM and MSS data covering very arid and dry environments, the Rayleigh relative scattering model seems to be appropriate for very clear conditions. Landsat TM calibrated natural color composite results generated using this model in desert environments confirms that it is acceptable for this type of conditions.

It should be emphasized here again that this is an improvement technique for the dark-object haze correction method and is not a technique by which the DN values for the ideal path radiance are computed for each band. The histogram method is used to identify the initial or starting haze value for one band, and then a relative scattering model, power law in this case, is used to predict the haze values for the other spectral bands. These values are then used to do a dark-

object type correction. The relative scattering model is used to predict the haze values for the spectral bands being used, given the haze value for one band. The relative scattering model is not used to compute the path radiance values from scratch.

The values of the power law functions used to represent the relative scattering models for Landsat TM and MSS spectral bands are shown in Table 1. The percent of scattering that each band contributes to the total scattering that occurs within all the bands is also shown for each model. The average wavelength values, in μm units, of the spectral bands of the Landsat TM and MSS systems were used to produce Table 1. Different units could have been used, but it is the relative relationship that is important; the spectral band values are ratioed and the units cancel out. The author realizes that the spectral band width of the individual bands affects the amount of radiance detected; however, the differences are not considered here because the bands affected the most by scattering, TM Bands 1, 2, 3 and MSS Bands 4 and 5 (visible part of the spectrum), have bandwidths that are very similar. From Table 1 it can be seen that using the relative scattering model of λ^{-4} for a very clear atmosphere TM Band 1 accounts for slightly over 50% of the total scattering and TM Band 5 for only 0.4% (sum of six nonthermal TM bands). This shows that the great majority of atmospheric scattering, which is an additive component, occurs in the visible part of the spectrum for a *very* clear atmosphere.

The relative scattering models presented here have generated good visual results. The computed haze values have been used to create TM calibrated natu-

TABLE 1 Values of Specific Functions for the Landsat TM and MSS Spectral Bands

	λ^a	VERY CLEAR λ^{-4} (%)		CLEAR λ^{-2} (%)		MODERATE λ^{-1} (%)		HAZY $\lambda^{-0.7}$ (%)		VERY HAZY $\lambda^{-0.5}$ (%)	
TM											
1	0.485	18.073	(50.5) ^b	4.251	(36.2)	2.062	(27.0)	1.659	(24.0)	1.436	(21.9)
2	0.560	10.168	(28.4)	3.189	(27.1)	1.786	(23.4)	1.501	(21.7)	1.336	(20.4)
3	0.660	5.270	(14.7)	2.296	(19.5)	1.515	(19.9)	1.338	(19.3)	1.231	(18.8)
4	0.830	2.107	(5.9)	1.452	(12.3)	1.205	(15.8)	1.139	(16.5)	1.098	(16.8)
5	1.650	0.135	(0.4)	0.367	(3.1)	0.606	(7.9)	0.704	(10.2)	0.778	(11.9)
7	2.215	0.042	(0.1)	0.204	(1.8)	0.451	(5.9)	0.573	(8.3)	0.672	(10.2)
Total	—	35.795		11.759		7.625		6.914		6.551	
MSS											
4	0.550	10.928	(52.2)	3.306	(38.6)	1.818	(31.7)	1.520	(29.2)	1.348	(28.3)
5	0.650	5.602	(26.8)	2.367	(27.7)	1.538	(26.8)	1.352	(26.0)	1.240	(26.0)
6	0.750	3.160	(15.1)	1.778	(20.8)	1.333	(23.2)	1.223	(23.5)	1.155	(24.2)
7	0.950	1.228	(5.9)	1.108	(12.9)	1.053	(18.3)	1.108	(21.3)	1.026	(21.5)
Total	—	20.918		8.559		5.742		5.203		4.769	

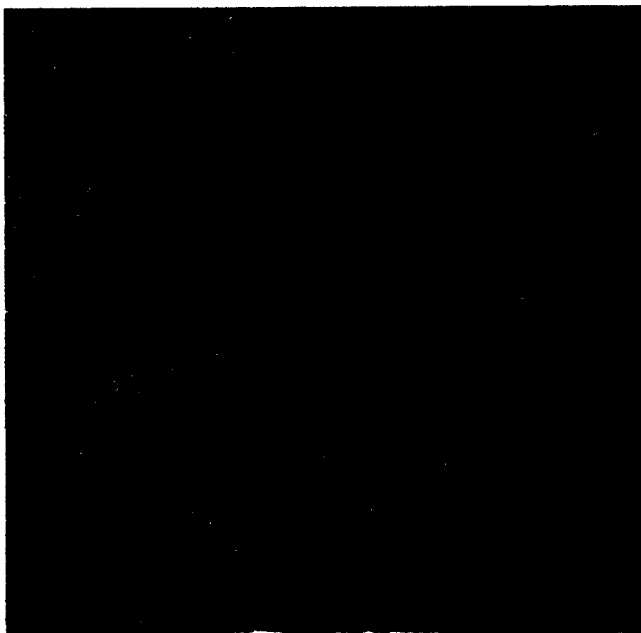
^aUsed average wavelength for the given spectral bands.

^bRepresents the percent of contribution of the given spectral band to the total scattering that occurs in all bands for the given model. This shows the *relative* percent of scattering that occurs in each band compared to the total scattering and is used to predict the scattering from the selected starting haze value. Remember that the starting haze value is selected using the histogram or dark-object subtraction technique. Also note that the starting haze value represents the atmospheric conditions during data collection (clear, hazy, etc.); therefore, the haze amplitude is in the proper range. Using these models for relative scattering, the TM spectral bands in the visible part of the spectrum (TM1, 2, and 3) account for 93.6%, 82.8%, 70.3%, 65.0%, and 61.1% of the scattering for very clear, clear, moderate, hazy, and very hazy atmospheres, respectively. Using these same models the MSS visible bands (4 and 5) account for 79.0%, 66.3%, 58.5%, 55.2%, and 54.3% percent of the scattering for very clear, clear, moderate, hazy and very hazy atmospheres, respectively. With the very hazy *relative* scattering model the bands become much more equal in the amount of scattering that occurs. This is because the scattering has become much more nonselective.

ral color composites using TM Bands 1, 2, and 3. Field investigations have confirmed the visual results seen on the final products. Figure 3 shows the results of two color composites made with Landsat TM Bands 1, 2, and 3. The composite in Fig. 3(a) shows the results of applying standard linear contrast stretches to the data. The composite in Fig. 3(b) shows the calibrated natural color results, which include the atmospheric correction discussed in this paper. The area lies in central Saudi Arabia, and the sands are highly oxidized, which accounts for their red color. Two color photographs are

shown in Figs. 3(c) and 3(d). They show the colors as they appear in the field, within the limitations of color reproduction in the photolab. The standard linear contrast stretch results are included only to show how the colors in this product compare to those in the calibrated results. It is understood that the standard linear contrast stretch results will often have higher contrast, which is desirable in some cases (e.g., visual interpretation). However, there are cases where working with calibrated data is more important (e.g., signature extension—spatial and temporal—and pattern recognition; this will

(a)



(b)

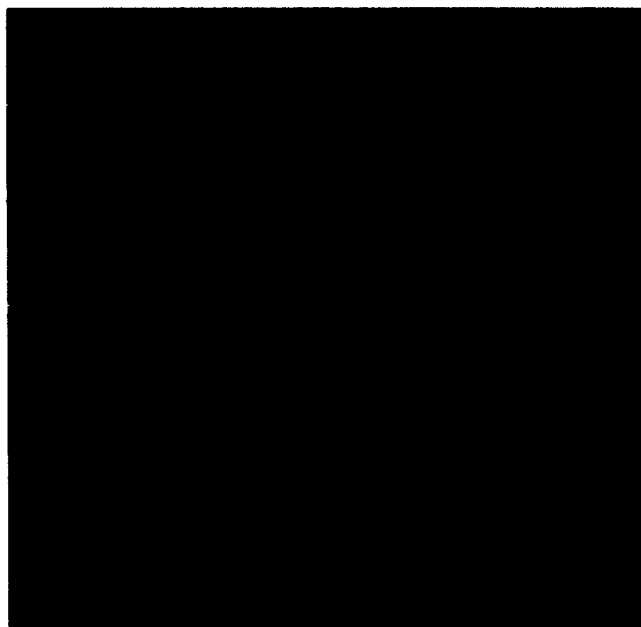


FIGURE 3. Thematic Mapper Bands 1, 2, and 3 color composite prints are shown in 3(a) and 3(b). The area is in central Saudi Arabia and the data were collected on 18 August 1984. The image data in 3(a) had standard linear contrast stretches applied and the image data in 3(b) were processed with the calibrated natural color algorithm, which includes the atmospheric correction discussed in this paper. Figures 3(c) and 3(d) are color photographs taken at two different sites in the field. The locations of these two sites are labeled c and d in 3(a) and 3(b). The view directions of the two photographs are approximately at a 45° angle towards the upper left. They show the colors as they appear in the field, within the limitations of color reproduction in the photolab (photographs courtesy of Graydon L. Berlin).

(c)



(d)



FIGURE 3. (Continued)

be especially important in the future when hundreds of spectral bands will be available and spectral signatures seen in the laboratory will be searched for in the image data).

A spectral band that is in the visible part of the spectrum should be used as the starting haze band. The user can select the starting haze value from any of the visible bands; however, he/she should use the one that shows more clearly the distinct increase in the number of pixels present at the lower end of the histogram. The starting haze DN value must not overpredict the values for the other bands; if this occurs, either a smaller starting haze value is required or selection of a starting haze DN value from another band may be necessary. This problem may be encountered in data that do not have a valid dark-object within the image, and this can be used to help identify and correct for this problem.

The starting haze value automatically has the proper amplitude for the given haze conditions, and analysis has indicated that it may be used as a guide to help select the relative scattering model that best represents the atmospheric conditions (i.e., very clear, clear, moderate, hazy, or very hazy). This can be done because the starting haze value selected with the histogram method represents the atmospheric conditions during data collection. Therefore, if the data were collected on a clear day, the starting haze value will be smaller than one collected during a hazy day. For example, a clear day's starting haze value for TM Band 1 might have a DN value of 55, while a hazy day's starting DN value might be equal to 90.

Table 2 shows the multiplication factors needed to compute or predict the

haze values for Landsat TM and MSS nonthermal bands when TM Band 1, 2, or 3 (or MSS Band 4 or 5) has been selected as the starting haze band. For example, assume that TM Band 1 is used as the starting band and a starting haze value of 40 is selected from its histogram [see Fig. 2(a)]; the predicted haze values for TM Bands 2, 3, 4, 5, and 7 for a very clear atmosphere will be 22.5, 11.7, 4.7, 3.0, and 0.1, respectively. These values were generated by multiplying the starting haze value of 40 by the very clear relative scattering model factors shown in Table 2 under TM Band 1 (i.e., 0.563, 0.292, 0.117, 0.075, and 0.002). If a hazy relative scattering model is selected, with TM Band 1 still used as the starting band, the predicted haze values for TM Bands 2, 3, 4, 5, and 7 would have been 36.2, 32.2, 27.5, 17.0, and 13.8, respectively. These values were generated using the hazy relative scattering model factors shown in Table 2 and are used only as an example. With a starting haze value of 40 the hazy model would not be applicable.

The haze values computed using the relative scattering models are not the correct DN values to remove the haze effects from the data. They must first be adjusted for the different gains and offsets used by the imaging system to collect the digital data. The DN values recorded by imaging systems can generally be represented by the following equation:

$$DN_i(X, Y) = GAIN_i * [Rad_i(X, Y)] + OFFSET_i, \quad (1)$$

where

$DN_i(X, Y)$ = output digital number for band i at pixel (X, Y) ,
 $GAIN_i$ = gain factor used for band i ,

TABLE 2 Multiplication Factors To Predict Haze Values in Other Spectral Bands
Given a Starting Haze Value (SHV) for Landsat 4 TM Bands 1, 2, or 3 and MSS 4 or 5^a

TM	VERY CLEAR λ^{-4}	CLEAR λ^{-2}	MODERATE λ^{-1}	HAZY $\lambda^{-0.7}$	VERY HAZY $\lambda^{-0.5}$
(TM1 = SHV)					
1 ^b	1.000 ^b	1.000	1.000	1.000	1.000
2	0.563	0.750	0.866	0.905	0.930
3	0.292	0.540	0.735	0.807	0.857
4	0.117	0.342	0.584	0.687	0.765
5	0.075	0.086	0.294	0.424	0.542
7	0.002	0.048	0.219	0.345	0.468
(TM2 = SHV)					
1	1.777	1.333	1.155	1.105	1.075
2 ^b	1.000	1.000	1.000	1.000	1.000
3	0.518	0.720	0.848	0.891	0.921
4	0.207	0.455	0.675	0.759	0.822
5	0.013	0.115	0.339	0.469	0.582
7	0.004	0.064	0.253	0.382	0.503
(TM3 = SHV)					
1	3.429	1.851	1.361	1.240	1.167
2	1.929	1.389	1.179	1.122	1.085
3 ^b	1.000	1.000	1.000	1.000	1.000
4	0.400	0.632	0.795	0.851	0.892
5	0.026	0.160	0.400	0.526	0.632
7	0.008	0.089	0.298	0.428	0.546
(MSS4 = SHV)					
4 ^b	1.000	1.000	1.000	1.000	1.000
5	0.513	0.716	0.846	0.889	0.920
6	0.289	0.538	0.733	0.805	0.857
7	0.112	0.335	0.579	0.729	0.761
(MSS5 = SHV)					
4	1.951	1.397	1.182	1.124	1.087
5 ^b	1.000	1.000	1.000	1.000	1.000
6	0.564	0.751	0.867	0.905	0.931
7	0.219	0.468	0.685	0.820	0.827

^aValues for five *relative* scattering models are shown and represent very clear, clear, moderate, hazy and very hazy scattering conditions.

^bMarks the band used to select the starting haze value. Entries to this table are computed by dividing the values in Table 1 for the given bands and relative scattering model by the Table 1 value of the selected starting haze band (e.g., if TM Band 2 is used to select the starting haze with a clear relative scattering model, then TM Bands 1 and 3 entries are computed by dividing 4.251 and 2.296 by 3.189, respectively; note that this is identical to dividing the relative percent of scattering that occurs in these bands).

$RAD_i(X, Y)$ = radiance value at pixel (X, Y) in band i ,

$OFFSET_i$ = offset factor used for band i .

The radiance value (RAD_i) is made up of multiplicative and additive terms. Some of the *major* contributors to the multiplicative term include the reflectance of the target at pixels (X, Y) in band i [$R_i(X, Y)$], the slope conditions at pixel (X, Y) [$SLOPE(X, Y)$], the sun elevation during data collection (SUN), and the atmospheric absorption characteristics in band i at the time the image is recorded (τ_i). The main contributor to the additive term is the haze present in band i [$HAZE_i$]. These factors can be used to obtain a general representation of the radiance parameter [$RAD_i(X, Y)$] in Eq. (1) as follows:

$$RAD_i(X, Y) = [R_i(X, Y) * SLOPE(X, Y) * SUN * (\tau_i)] + HAZE_i. \quad (2)$$

If the multiplicative factors are combined into a single variable called $MULT_i(X, Y)$, and we substitute Eq. (2) into Eq. (1), the following simple relationship is generated:

$$DN_i(X, Y) = GAIN_i * [MULT_i(X, Y) + HAZE_i] + OFFSET_i. \quad (3)$$

From this equation it is easy to see that in order to correctly predict the additive scattering or haze values for the different spectral bands in an imaging system, using a starting haze value from one of the bands and a relative scattering model, the

data must be corrected or normalized for the different gain and offset parameters used. Table 3 lists the gain and offset parameters used by the Landsat 4 TM and MSS imaging systems. Parameters used by the Landsat 1, 2, and 3 MSS systems are given by Robinove et al. (1981).

Using Eq. (1) to solve for the radiance, the following simple relationship results:

$$RAD_i(X, Y) = \frac{DN_i(X, Y) - OFFSET_i}{GAIN_i}. \quad (4)$$

This relationship can be used to correct the data for gain and offset effects. It converts the data from digital counts to radiance values ($mW cm^{-2} sr^{-1} \mu m^{-1}$), which represent standard physical units. This type of conversion is desirable under certain conditions, as discussed by Robinove (1982). However, if desired, the predicted haze values can have the proper gain and offset effects introduced so that they have the proper DN value. This can be done by generating gain normalization factors that can be applied to the predicted haze values along with the offset corrections to predict the DN that the desired haze value would be mapped to in the given band. The normalization factors needed for Landsat 4 TM and MSS data are shown in Table 3. They were produced by dividing the starting haze TM band (or MSS band) gain value into the gain value of each particular spectral band.

These normalization values, along with the offsets, allow the predicted haze values to be mapped into the proper DN range for their corresponding bands. This

TABLE 3 Gain, Offset, and Gain Normalization Factors Used for Landsat-4 TM and MSS Images

TM	GAIN	OFFSET	NORMALIZATION ^a		
			TM1	TM2	TM3
1	15.78	2.58	1.00	1.95	1.49
2	8.10	2.44	0.51	1.00	0.76
3	10.62	1.58	0.67	1.31	1.00
4	10.90	1.91	0.69	1.35	1.03
5	77.24	3.02	4.89	9.54	7.27
7	147.12	2.41	9.32	18.16	13.85

MSS	GAIN	OFFSET	NORMALIZATION	
			MSS4	MSS5
4	55.70	− 1.11	1.00	0.77
5	72.16	− 2.89	1.29	1.00
6	100.79	− 4.03	1.81	1.40
7	15.91	− 0.16	0.29	0.22

^aNormalization factors are computed by dividing the starting haze TM band (or MSS band) gain value into the gain values of the other given bands (e.g., TM Band 2 entry of 0.51 was generated by dividing 8.10 by 15.78). TM gain and offset values from Barker (1985, p. 297) and MSS values computed from specified RMIN and RMAX values given by Alford and Imhoff (1985), and assuming a maximum DN of 127 for MSS Bands 4, 5, and 6, and 63 for Band 7. Gains are in counts per mW cm^{−2} sr^{−1} μm^{−1} and offsets are in DN counts.

can be done using Eq. (3), where the multiplicative term (MULT_{*i*}) is assumed to be zero because *R_i* is zero, and the gain factor (GAIN_{*i*}) is normalized as shown in Table 3 (NORM_{*i*}). This gives the following relationship:

DN(HAZE_{*i*}) = NORM_{*i*} * HAZE_{*i*}

+ OFFSET_{*i*}, (5)

where

- DN(HAZE_{*i*}) = DN value that predict-
- ed HAZE_{*i*} value will be
- mapped to in band *i*,
- NORM_{*i*} = normalized gain value
- from Table 3,
- HAZE_{*i*} = predicted haze value for
- band *i* using a relative
- scattering model and

the offset corrected

starting haze value,

OFFSET_{*i*} = offset value used for

band *i* from Table 3.

For example, the minimum dark-object DN value for TM Band 1 for the image shown in Fig. 1 is equal to 40, and when corrected for its offset value of 2.58, the TM band 1 starting haze value is equal to 37.42. Using the very clear relative scattering model multiplication factors for TM band 1 (Table 2), the predicted haze values for TM Bands 2, 3, 4, 5, and 7 are 21.1, 10.9, 4.4, 2.8, and 0.1, respectively. These predicted HAZE_{*i*} values, along with the normalization values (NORM_{*i*}) shown in Table 3, can be used in Eq. (5) to obtain the DN values that the desired/predicted haze values will be mapped to after the corresponding gain and offset factors have been applied. Using this method, the DN value for TM Band 2

can be generated as follows:

$$\begin{aligned}\text{HAZE}_2 &= (\text{TM 1 starting here}) * (\text{TM2} \\ &\quad \text{relative scattering model multi-} \\ &\quad \text{plication factor from Table 2}) \\ &= (37.42) * (0.563) \\ &= 21.067.\end{aligned}$$

Then, using Eq. (5),

$$\begin{aligned}\text{DN}(\text{HAZE}_2) &= (\text{NORM}_2 \text{ from Table} \\ &\quad 3) * (\text{HAZE}_2) + (\text{OFFSET}_2 \\ &\quad \text{from Table 3}) \\ &= (0.51) * (21.067) + (2.44) \\ &= 13.184.\end{aligned}$$

The DN values that represent the predicted haze values for TM Bands 2, 3, 4, 5, and 7 are 13.2, 8.9, 4.9, 16.7, and 3.3, respectively. Notice the dramatic change in the relationship between the haze values that the different gains and offsets have made. It can be seen from these values, which are also shown in Table 4, that the original dark-object DN values selected with the old method were close, but not exactly the same, as the predicted values. In this particular case, the very steep-walled canyon of the Little Colorado River gorge (Fig. 1) produced a

TABLE 4 DN Values Generated for Haze Corrections Using the Old/Standard and Improved Dark-Object Subtraction Techniques for One Landsat TM and Two MSS Images^a

A. LANDSAT TM IMAGE COVERING THE NORTHERN ARIZONA AREA SHOWN IN FIG. 1			
BAND NO.	OLD ^b	PREDICTED ^c	FINAL ^d
1	40.0	37.4	40.0
2	13.0	21.1	13.2
3	12.0	10.9	8.9
4	8.0	4.4	4.9
5	5.0	2.8	16.7
7	2.0	0.1	3.3
B. Landsat 1 MSS Image Covering an Area in Western Tunisia			
4	13.0	13.0	13.0
5	9.0	6.7	8.6
6	21.0	3.8	6.9
7	16.0	1.5	0.4
C. Landsat 2 MSS Image Covering an Area in Southern Tunisia			
4	30.0	30.0	30.0
5	50.0	26.7	34.4
6	50.0	24.1	43.6
7	40.0	21.9	6.4

^aThe amplitude of the starting haze values for the TM and first MSS images indicates the use of the very clear relative scattering model; the amplitude of the starting haze value of the second MSS image indicates the use of the hazy relative scattering model.

^bHaze values using old/standard dark object subtraction method.

^cPredicted haze values, using TM Band 1 or MSS Band 4 as the starting haze value and the appropriate relative scattering model.

^dPredicted haze values corrected for the gain and offset values used by each band. These are equal to the actual DN values to be subtracted from the given spectral bands. Notice that because of the gain differences it is possible for the final haze value for a longer wavelength band (MSS Band 6) to have a larger value than a shorter wavelength band (MSS Band 5).

very good dark-object or shadowed area, and so the old method was close. However, images are often encountered where this is not the case, and a good dark-object does not exist. In such images the old method can generate very different values from the predicted ones generated with the improved method.

Table 4 shows the haze values for the TM image just discussed and two Landsat MSS images using the old and improved methods. One of the Landsat MSS images, which covers an arid and vegetated region, has haze values for Bands 4, 5, 6, and 7 of 13, 9, 21, and 16 using the old method and predicted haze values of 13, 6.7, 3.8, and 1.5 using the improved method presented here. The very clear relative scattering model was used due to the low MSS Band 4 starting haze value. A second MSS scene, covering an arid region with little topography, has haze values of 30, 50, 50, and 40 using the old method and values of 30, 27.6, 24.1, and 21.9 using the new method. The hazy relative scattering model was used due to the high MSS Band 4 starting haze value. After gain and offset normalization, the predicted haze values for the two MSS images are mapped to DN values of 13, 8.6, 6.9, 0.4 and 30, 34.4, 43.6 and 6.4, respectively. Notice that due to the gain differences the DN values to be used for haze correction for longer wavelength bands can, at times, be larger than the shorter wavelength bands. The differences between the old and improved methods for these two scenes are quite dramatic and show that if the old dark-object subtraction method is used without consideration to the relative scattering that occurs in the atmosphere, the results can generate unrealistic haze correction values.

There is a problem with the haze DN values generated for TM Bands 5 and 7. The predicted haze values look acceptable, but the final DNs derived using Eq. (5) are too high. If these values are used, it will map several DN levels of good data to zero, which would be an overcorrection. The reason why this happens is not clear, but it may be related to the extreme difference in the gain values between these two bands and the first four (see Table 3). This in turn is related to the fact that the maximum radiance level seen by these two bands is much lower than the others, which may influence the DN or bit range (Alford and Imhoff, 1985). Part of the problem may be due to the band width differences (TM Bands 1, 2, and 3 vs. TM Bands 5 and 7) and indicates that in some cases it is important to consider its effects. However, the amount of atmospheric scattering that occurs at this region of the spectrum is quite small, except for very hazy atmospheres and can be considered negligible, and a haze DN value of zero, or one DN at the most, can be used for Bands 5 and 7.

Table 5 shows the corresponding haze values to be used for Landsat 4 TM Bands 1, 2, 3, and 4 data using TM Band 1 as the starting haze value. The table was generated using the five power law relative scattering models discussed in this paper, and a similar table can be generated using other models and for other multispectral imaging systems. Table 5 is set up to use the appropriate relative scattering model based on the amplitude of the starting haze value of TM Band 1 (i.e., use the very clear relative scattering model for small starting haze values up to the very hazy relative scattering model for very large starting haze values). A

TABLE 5.^a

RELATIVE SCATTERING MODEL	TM1 ^b	TM2	TM3	TM4
Very clear (≤ 55)	30.0	11.1	7.4	4.3
	35.0	12.5	8.4	4.7
	40.0	13.9	9.4	5.1
	45.0	15.4	10.4	5.5
	50.0	16.8	11.4	5.9
	55.0	18.2	12.3	6.4
Clear (56–75)	60.0	25.4	23.3	16.1
	65.0	27.3	25.1	17.2
	70.0	29.2	26.9	18.4
	75.0	31.1	28.7	19.6
Moderate (76–95)	80.0	37.8	41.0	34.1
	85.0	40.0	43.4	36.2
	90.0	42.2	45.9	38.2
	95.0	44.4	48.4	40.2
Hazy (96–115)	100.0	48.6	55.6	49.3
	105.0	50.9	58.4	51.7
	110.0	53.2	61.1	54.1
	115.0	55.5	63.8	56.4
Very hazy (> 115)	120.0	59.4	70.5	65.3
	125.0	61.7	73.4	67.9
	130.0	64.1	76.2	70.5
	135.0	66.5	79.1	73.2

^aThe DNs shown above are equal to the predicted haze values after gain and offset normalization for Landsat 4 TM Bands 1, 2, 3, and 4. The functions of λ^{-4} , λ^{-2} , λ^{-1} , $\lambda^{-0.7}$, and $\lambda^{-0.5}$ were used to represent the relative scattering of very clear, clear, moderate, hazy, and very hazy atmospheric conditions, respectively. Due to the gain differences between the bands the DN value used to correct for haze of a longer wavelength band can be larger than one used for a shorter wavelength band (e.g., TM Bands 2 and 3). The difference in gains also makes it difficult to see the relative scattering relationship (e.g., TM Band 1 starting haze value of 100 predicts a TM Band 2 haze value of 90.5; however, 48.6 must be used due to the difference in gain).

^bTM Band 1 was used as the starting haze value band. Notice the jump that occurs between the functions used. This could be eliminated by using a continually variable power model rather than just five discrete values. The amplitude of the starting haze value (SHV) is used to determine the type of atmospheric conditions that existed during data collection. The relationship used in this table was derived from analysis of over 25 TM scenes. The entries to the table were computed as follows:

$$\text{DN(OUT)} = (\text{SHV of TM Band 1}) * (\text{Table 2 factor}) * (\text{Table 3 NORM}) \\ + (\text{Table 3 OFFSET}).$$

table could be generated where the power used in the relative scattering model is continually variable rather than just five discrete values, as in this particular case. This would allow the relative scattering function to change continuously and eliminate abrupt boundary changes.

Summary

A method has been developed that allows atmospheric scattering parameters to be generated so that their DN values conform to a user-selected *relative* scattering model. The computed haze values, which are wavelength-dependent and correlated to each other, can be used after proper gain and offset normalization to apply a simple dark-object subtraction to remotely sensed multispectral image data to correct for atmospheric scattering. In this study the relative scattering models used were power law models where the power used was based on the amplitude of the starting haze value. Other possible relative scattering models could be used. The critical aspect of this technique is to use DN values that conform to some realistic relative scattering model so that the haze values will be wavelength-dependent and correlated with each other.

The haze corrections can be applied to the original data in DN counts of the imaging system. However, the user can, if desired, convert the entire image to radiance units before correcting for haze. But normalizing the predicted haze values for gain and offset allows the corrections to be applied without having to convert an entire image's DNs into radiance values. This does not imply that it is better to use the DN counts, but that an optional way exists to accomplish the improved correc-

tion without having to convert the entire image to radiance values.

This type of correction is important if remotely sensed data are to be exploited to their maximum potential. Studies dealing with analyses of the spectral response of cover types, temporal studies, and signature extension (temporally and spatially) have to be able to correct or remove variable factors, such as atmospheric scattering, from the data which have nothing to do with the information of interest. Research is currently under way to establish a better relationship between the starting haze value and the type of atmospheric conditions present during data collection.

References

- Ahern, F. J., Goodenough, D. G., Jain, S. C., Rao, V. R., and Rochon, G. (1977), Landsat atmospheric corrections at CCRS, In *Proceedings of the Fourth Canadian Symposium on Remote Sensing*, Quebec City, May, pp. 583–595.
- Ahern, F. J., Bennett, D. M., Guertin, F. E., Thomson, K. P. B., and Fedosejeus, G. (1982), An automated method for producing reflectance—enhanced Landsat images, In *Proceedings of 1982 Machine Processing of Remotely Sensed Data Symposium*, Purdue University, pp. 328–336.
- Alford, W. L., and Imhoff, M. L. (1985), Radiometric accuracy assessment of Landsat-4 Multispectral Scanner data, *NASA Conference Publication 2355, Early Landsat-4 Results* I:9–21.
- Barker, J. L., Ball, D. L., Leung, K. C., and Walker, J. A. (1985), Prelaunch absolute radiometric calibration of the reflective bands on the Landsat-4 protoflight Thematic Mapper, *NASA Conference Publication 2355, Early Landsat-4 Results* II:277–373.

- Carr, J. R., Glass, C. E., and Schowengerdt, R. A. (1983), Signature extension versus retraining for multispectral classification of surface mines in arid regions, *Photogramm. Eng. Remote Sens.* 49(8):1193–1199.
- Chavez, P. S., Jr. (1975), Atmospheric, solar, and MTF corrections for ERTS digital imagery, In *Proceedings of the American Society of Photogrammetry*, Fall Technical Meeting, Phoenix, AZ, p. 69.
- Chavez, P. S., Jr., Berlin, G. L., and Mitchell, W. B. (1977), Computer enhancement techniques of Landsat MSS digital images for land use/cover assessments, In *Proceedings of the Sixth Remote Sensing of Earth Resources Symposium*, Tullahoma, TN, pp. 259–275.
- Curcio, J. A. (1961), Evaluation of atmospheric aerosol particle size distribution from scattering measurement in the visible and infrared, *J. Opt. Soc. Am.* 51:548–551.
- Holben, B., and Justice, C. (1981), An examination of spectral band ratioing to reduce the topographic effects on remotely sensed data, *Int. J. Remote Sens.* 2(2): 115–133.
- Otterman, J., and Robinove, C. J. (1981), Effects of the atmosphere on the detection of surface changes from Landsat Multispectral Scanner data, *Int. J. Remote Sens.* 2(4):351–360.
- Robinove, C. J. (1982), Computation with physical values from Landsat digital data, *Photogramm. Eng. Remote Sens.* 48(5): 781–784.
- Robinove, C. J., Chavez, P. S., Jr., Gehring, D., and Holmgren, R. (1981), Arid land monitoring using Landsat albedo difference images, *Remote Sens. Environ.* 11:133–156.
- Rowan, L. C., Wetlaufer, P. H., Goetz, A. F. H., Billingsley, F. C., and Stewart, J. H. (1974), Discrimination of rock types and detection of hydrothermally altered areas in south-central Nevada by the use of computer-enhanced ERTS images, *U.S. Geological Survey Professional Paper* 883, 35 pp.
- Sabins, F. F., Jr. (1978), *Remote Sensing Principles and Interpretation*, Freeman, San Francisco, p. 18.
- Siegal, B. S., Gillespie, A. R., and Skaley, J. E. (1980), *Remote Sensing in Geology*, Wiley, New York, p. 120.
- Slater, P. N., Doyle, F. J., Fritz, N. L., and Welch, R. (1983), Photographic systems for remote sensing, *American Society of Photogrammetry Second Edition of Manual of Remote Sensing*, Vol. 1, Chap. 6, pp. 231–291.
- Spanner, M. A., Peterson, D. L., Hall, M. J., Wrigley, R. C., Card, D. H., and Running, S. W. (1984), Atmospheric effects on the remote sensing estimation of forest leaf area index, In *Proceedings of Eighteenth International Symposium on Remote Sensing of Environment*, Paris, France, pp. 1295–1308.
- Turner, R. E., Malila, W. A., and Nalepha, R. F. (1971), Importance of atmospheric scattering in remote sensing, In *Proceedings of Seventh International Symposium on Remote Sensing of Environment*, Ann Arbor, MI, pp. 1651–1697.
- Vincent, R. K. (1972), An ERTS Multispectral Scanner experiment for mapping iron compounds, In *Proceedings of the Eighth International Symposium on Remote Sensing of Environment*, Ann Arbor, MI, pp. 1239–1247.
- Vincent, R. K. (1973), Spectral ratio imaging methods for geological remote sensing from aircraft and satellites, In *Proceedings of the American Society of Photogrammetry, Management and Utilization of Remote Sensing Data Conference*, Sioux Falls, SD, pp. 377–397.