An Image Transform to Characterize and Compensate for Spatial Variations in Thin Cloud Contamination of Landsat Images

Y. Zhang[1, 2], B. Guindon[1] and J. Cihlar[1]

[1]Canada Centre for Remote Sensing
588 Booth Street, Ottawa
Canada, K1A 0Y7

[2]Intermap Technologies, 2 Gurdwara Road
Ottawa, Canada, K2E 1A2

Corresponding Author:
Ying Zhang, Dr.
588 Booth Street, Ottawa
Canada, K1A 0Y7
Tel. (613) 943-2006
Fax. (613) 947-1383
ying.zhang@ccrs.nrcan.gc.ca

Abstract

A Haze Optimized Transformation (HOT) is developed and assessed for the detection and characterization of haze/cloud spatial distributions in Landsat scenes. The transformation is derived from an analysis of a visible-band space where spectral response to diverse surface cover classes under clear sky conditions is highly correlated but spectral response to haze is highly sensitive to both optical wavelength and haze optical depth. The robustness of the detection algorithm is demonstrated through its application to visible band imagery of seven Landsat 5 Thematic Mapper (TM) and Landsat 7 Enhanced Thematic Mapper Plus (ETM+) scenes that encompass diverse surface cover and atmospheric characteristics. A methodology for utilizing the transformed image to radiometrically compensate visible band imagery is presented and quantitatively tested in the correction of an example ETM+ scene.

1. Introduction

Users of satellite scenes for terrestrial applications are faced with two important realities. First, atmospherically clear scenes are seldom available for the temporal window of interest even in the case of the restricted geographic coverage of Landsat scenes (e.g. 185 km x 185 km). Second for hazy scenes, there typically is a paucity of ancillary data upon which to base an absolute atmospheric correction. Contamination by spatially varying, semi-transparent cloud and aerosol layers (hereafter referred to by the generic term 'haze') is a common problem that affects a significant portion of scenes within existing Landsat archives. Haze can arise from a variety of atmospheric constituents including water droplets, ice crystals or fog\smog particles (Kaufman, 1989). Its resulting influence on measured radiance varies significantly among the six 30-metre resolution TM and ETM+ spectral bands (note: hereafter we will use the term 'TM' to refer to the 30-metre bands of both sensors), being most pronounced within the visible spectral region. There remains a pressing need for a robust, image-based capability to accurately detect and isolate spatially-varying forms of haze and to compensate for its influence on radiance in at least a relative (i.e. within-scene normalization) sense. Such an ability would have a significant impact on many applications, in particular land cover classification and mapping (e.g. Song et al., 2001).

A number of studies on image-based atmospheric correction have been undertaken and here we briefly summarize the ability of the most common approaches to model spatial variations in haze. First, dark target subtraction methods for haze removal have been extensively developed and utilized (Chavez 1988; Chavez, 1989; Teillet and Fedosejevs, 1995; Teillet et al. 1987). In general the approach involves the estimation of band-specific grey level offsets from histogram lower bounds. These offsets are then applied scene-wide. In principal, the method can be extended to model intra-scene haze variations by partitioning a scene into sub-areas and estimating histogram lower bounds for each (e.g. Teillet et al., 1987). However, since accurate histogram lower bound estimates require large pixel populations, this extension can only be used to model coarse-scale haze structure (i.e. on the scale of tens of kilometers). Liang et al. (1997) have proposed an alternate strategy that combines 'dark object' detection and physical atmospheric modelling. This technique has the advantage that isolated pixels (e.g. dense vegetation) are sought thereby allowing for the characterization of more detailed haze structure (i.e. down to the scale of a kilometer).

A second approach involves isolating the haze contribution in a scene through a radiometric transformation. During the extension of the Tasselled Cap (TC) transform (Kauth and Thomas, 1976) to the Thematic Mapper (TM) sensor, it was noted that haze seemed to be a major contributor to the 4th TC component (Crist et al. 1986; Crist and Cicone, 1984a; Crist and Cicone, 1984b; Lavreau, 1991; Richter, 1996a; Richter, 1996b). It is important to note that the original TC transform was tailored to separate surface radiometric contributions of soil and vegetation (i.e. 'brightness' and 'greenness' components) and hence is not optimized for haze detection. Subsequently, Richter (1996a) developed an aerosol haze removal methodology based on a reduced two-band version (TM bands one and three) of this Tasselled Cap transform. Richter's approach required complex procedures to compensate for the surface cover information also captured by the transform. In addition, he applied it to a restricted data set composed of a single TM subscene.

The work reported here has resulted from the need for a practical, scene-based, simple-to-use methodology for rudimentary haze removal from Landsat imagery in order to improve land cover mapping performance. Since a relative intra-scene normalization is sought that is capable of detecting and accounting for haze with fine-scale structure, we have elected an image transformation approach.

This paper reports on results of this study and addresses two objectives:

(a) To develop a robust transformation to characterize the spatial distribution of the contribution of haze and clouds to image radiometry. This transformation has been designed to minimize terrestrial surface influences and be applicable to imagery from all Landsat sensors. Extensive use is made of atmospheric radiation transfer modelling to support this development.

(b) To present a methodology that employs the haze transformed image to radiometrically adjust visible band imagery. As a 'proof of concept', an example TM scene is processed and quantitatively assessed.

A new ‘haze optimized transformation’, hereafter referred to as HOT, is introduced and developed in section 2. The results from a modelling investigation of the sensitivities of HOT to various atmospheric conditions are described in section 3. In section 4, haze distributions, extracted from seven Landsat TM images using the transformation are presented and the capability of the transformation to extract the haze spatial variation is examined and compared with the fourth component of the Tasselled-Cap transformation. Section 5 outlines a methodology to utilize the HOT image to radiometrically adjust visible band imagery. An assessment of the HOT-based method using paired hazy and adjusted images is also described in this section.

2. Haze Optimized Transformation (HOT)

An optimized transformation will be one that quantifies atmospheric influences on satellite image radiometry but at the same time is insensitive to surface reflection effects. A number of criteria must be met to achieve these goals.

(a) A spectral space (i.e. a mix of spectral bands) must be selected within which the spectral responses of different land cover types, under clear atmospheric conditions, are highly correlated. This will result in a well-defined surface response vector in spectral space. Hereafter we will refer to this vector as the ‘clear line’ (CL).

(b) The effect of haze on apparent radiance must be different for the component bands of the selected spectral space. As a result, increased atmospheric contamination will manifest itself in increased migration away from the CL.

(c) A suitable transformation will be one whose coefficients define a direction orthogonal to the CL and whose response magnitude is proportional to the deviation from this line. The conceptual framework of this approach is similar to that of Kauth and Thomas (1976) in which orthogonalization was employed to separate surface contributions of soil and vegetation cover. The difference here however is that the spectral responses to surface cover and atmospheric variations in spectral space are not orthogonal and therefore the transform response will only be proportional to atmospherically-induced radiance.

A simple spectral space, consisting of two visible bands, meets the first two criteria, for example, bands one and three in the case of the Landsat TM (hereafter TM1 and TM3). To validate this selection, we have generated clear sky radiance estimates for TM1 and TM3 for a set of representative surface cover types of Canadian landscapes (Table 1). The surface reflectance spectra were retrieved from 4-metre resolution PROBE-1 measurements (Secker et al., 2001) and used as input for MODTRAN, an atmospheric radiation transfer model (Berk et al. 1999). PROBE-1 samples were averaged to simulate the 30-metre resolution of the Thematic Mapper. Figure 1 illustrates the position of each surface class, coded as A to K, in the TM1-TM3 spectral space. High spectral correlation is apparent, resulting in a well-defined CL. The correlation coefficient for the clear-sky points A to K is 0.993. In this simulation it was assumed that the top-of-the-atmosphere (TOA) radiance for a surface type does not include the radiation scattered by the surroundings, i.e. the so-called 'adjacency effect' was ignored. In a real image, adjacency effects would generate a range in the pixel radiance within a surface type around the 'ideal' value. Therefore, the correlation coefficient for real images can be expected to be less than 0.99 but still be significantly high. The direction of the CL can be expressed by its slope angle, Θ, and hence HOT, the transformation that quantifies the perpendicular displacement of a pixel from this line will be given by;

$$\mathrm{HOT} = B_1 \sin\Theta - B_3 \cos\Theta \qquad (1)$$

where B_1 and B_3 are the pixel's band one and band three digital numbers (DNs) respectively.

In practice the proposed method can only be applied in a relative rather than absolute sense since absolute information regarding the atmospheric conditions over most scenes will not be available. Therefore, the angle Θ must be estimated from pixels selected from areas of a scene that visually are deemed to be the clearest, and hence Θ will vary somewhat scene-to-scene. HOT will then be employed to adjust the rest of the scene to the same atmospheric conditions as these clearest areas, i.e. to undertake a relative, not absolute intra-scene balancing.

To better understand the effects of atmospheric contamination in the TM1-TM3 spectral space, MODTRAN version 4 code was employed to estimate apparent TM1 and TM3 radiances for nineteen different levels of atmospheric optical depth. Atmospheric profiles for mid-latitude-summer conditions and haze contamination arising from thin-layered stratus cloud were studied. The optical depth of this cloud at 0.55 μm has been varied between 0 and 6.7 in eighteen equal increments. A solar zenith angle of thirty-eight degrees was assumed in these estimates although the ratio of band radiances and hence the clear line slope should be independent of this parameter. Rayleigh scattering and gas absorption of the idealized standard atmosphere are included in the calculation. The Landsat-5 sensor gains and offsets were employed during the conversion from radiances to digital numbers (DNs). The simulated radiances were calculated for an ideal atmosphere without the background aerosols since the background aerosol effect is linearly related to other path scattering effects. If the scattering effect of the background aerosols were to be included, the position of the CL would shift but its slope would not change appreciably.

Figure 1 shows the resulting 'haze trajectories' for each surface class, illustrating the migration away from the CL with increasing contamination (increasing optical depth). It should be noted that the slopes of lines of constant cloud optical depth are similar to the slope of the true CL. For example the slope of clear-sky CL in Figure 1 is 1.03 while the slope of the cloud line at level nineteen is 0.98. This implies that a first order radiance adjustment for haze can be achieved for TM1 and TM3 using suitably scaled HOT values computed from the image data. On the other hand, a closer inspection of Figure 1 reveals that the slopes of the haze trajectories monotonically decrease for surface classes of increasing absolute reflectance in these bands. This means that the proposed HOT is not completely insensitive to surface reflectance and may require, as a second-order term, a surface-dependent adjustment.

An important aspect of a transformation for haze detection must be its insensitivity to surface reflectance variations. For the 12 surface classes listed in Table 1, we have estimated clear sky radiances for all 30-metre resolution TM bands. From these we have estimated corresponding response levels of the HOT and 4th Tasselled Cap (TC) transformations (Crist and Ciccone, 1984b). Figures 2(a) and 2(b) illustrate plots of HOT (uppercase letters) and TC (lower case letters) values vs. DN level for TM1 and the near-infrared band TM4 respectively. The latter band is included to illustrate the potential of applying HOT to bands beyond the visible portion of the spectrum. Qualitatively, the HOT values of different surface cover classes are similar for the same haze level, and, hence, to first order, this transformation is independent of cover class. On the other hand, the TC values exhibit significant variation, indicating that the TC transformation is much more sensitive to surface reflectance. Further modelling (not shown) has indicated that TC response tends to increase both as a result of increasing cloud optical depth and increasing surface reflectance. This implies a surface-induced increase can be mistaken for haze in the TC transform image. Bright land surface types (urban, road and bare soil) are the most confusing since they trigger TC responses comparable in magnitude to those observed in high haze conditions over vegetated terrain. In Richter’s (1996b) methodology, based on a TC hybrid, additional steps were needed to identify and exclude urban regions.

Finally, in a practical application of transformations such as the HOT or TC, a threshold must be specified below which no haze compensation is attempted, assuming that lower transformed values are an unresolved mixture of haze and surface components. To exclude the surface effects, the haze-equivalent threshold of TC must be much higher than that of HOT. To examine these effects, scatter plots were generated of HOT (Figure 3a and 4a) and TC (Figure 3b and 4b) vs. TM1 and TM4 respectively. The plots were derived for a clear-sky urban/vegetation sub-image of a Landsat-5 TM scene of P21/R26 acquired on August 12, 1998. The observed tendency and change behaviour in the HOT-TM space for both TM bands (Figures 3a and 4a) agree well with model predictions. This confirms, with real imagery, the earlier model prediction of the superior performance of the HOT in the minimization of surface effects. It also provides evidence that simulation can capture the essential features of real data behaviour and therefore offers a powerful tool to aid in the design and development of practical, image-based methodologies.

3. Sensitivity to Cloud and Aerosol Types

In this subsection we address the issue of the dependence of HOT response to differing types of haze. To accomplish this, one surface class has been selected (conifer forest) and haze trajectories have been generated for it for various cloud and aerosol conditions (Figure 5). Besides stratus clouds (Figure 1), cirrus clouds consisting of ice crystals at high altitudes, rural aerosols with 23- and 5-km visibility (type 2 and 3), urban aerosols with 5-km visibility (type 4) and a background aerosol type with 50-km visibility (type 5) were modelled. For comparison, an ideal atmosphere without cloud and aerosols (type 1) is also presented. To generate the appropriate trajectories, ice cloud optical depth at 0.55 μm was varied from 0 to 3.6, while water cloud optical depth ranged from 0 to 2.8. The HOT values increased with increasing cloud optical depth. For four types of atmospheric particulates (ice crystals, water droplets, urban aerosols and rural aerosols) with the same optical depth at 0.55 μm ($\delta_{0.55} = 1.27$), the corresponding radiances in each TM band varied due to differences in the scattering and absorption properties (Figure 6). Summarizing Figures 5 and 6, the apparent radiances in the TM visible bands increased with the atmospheric condition in the order ice cloud, urban aerosols, water cloud and rural aerosols, with a same $\delta_{0.55}$, as did the HOT response. Ice clouds reflect less than water clouds in all of TM bands, thus, ice clouds appear darker in TM images than water clouds with the same optical depth. The HOT values of ice clouds were also lower than those of water clouds with the same cloud optical depth. In addition, the model results also show that TM1 and TM3 DNs increased but TM4 remained constant with increases in HOT response in the low optical depth range of clouds and aerosols. This implies that for optically very thin cloud or aerosol layers over coniferous forests at least, there should be no need to radiometrically compensate TM4 data.

Since the absorption and scattering properties of atmospheric aerosols can vary greatly depending on the aerosol composition, we expect that different aerosols, exhibiting the same optical depth, will not induce the same spectral radiance change in a given TM band. For example, urban aerosols (type 5) exhibit higher absorption in TM4 due to their higher soot content (WMO, 1986). For urban areas with heavy air pollution, we expect that TM4 DN level will decrease with increasing HOT response. On the other hand, for a rural aerosol layer, the change in DN level will be less in TM4 than in the TM visible bands as HOT increases. In summary, the simulation results indicate that the relationship between HOT values and TM radiances varies with different cloud/aerosol types. Nevertheless, the haze trajectories for different types of haze are near-coincident especially in the visible bands (Figure 5a and 5b), i.e. their dispersion is much lower than the radiometric dynamic range associated with the spectrum of surface cover types. This implies that knowledge of haze type is not needed to undertake a first order HOT-based radiometric compensation. On the other hand, the HOT value cannot be used to infer optical depth without additional information.

4. Example Results

To test the effectiveness of HOT, we applied it to seven scenes covering representative Canadian landscapes and encompassing a broad spectrum of atmospheric conditions ranging from clear conditions through translucent haze to opaque cloud. A summary of the scenes and their clear line descriptors (i.e. slopes and correlation coefficients) are presented in Tables 2 and 3. HOT (left panel) and TM3 (right panel) image pairs for six of the seven cases are illustrated in Figures 7. Although the CL slopes exhibited some variation, all scenes exhibited high (≥0.87) correlation coefficients.

The haze/cloud types in these seven images include both 'natural' (i.e. cumulus,status and cirrus) and 'man-made' (i.e. aircraft contrails) conditions. The surface types in these images cover a broad spectrum of visible-band reflectance from clear water at the low end to urban and snow/ice at the high end. The most common surface features in all scenes are forests and inland water bodies.

In scene P21/R26-a (Figure 7a), a contrail-induced middle-level system, created by the shearing effect of cross winds, occupies the left portion of the image. A contrail at about 9-10 km altitude (estimated from solar zenith angle and the distance between the contrail and its shadow) is above it. A thin haze layer covers a large area in the left part of the same image. The right part of the image appears the clearest and this region was used as the clear-sky area to define the HOT 'clear line'. The terrestrial surface cover is primarily a mix of forest/vegetation types with some urban/road features in the upper portion of the image.

A second image of P21/R26, hereafter referred to as P21/R26-b, has also been studied. While it would be considered clear overall, based on a cursory inspection, the HOT image and a more detailed visual analysis reveals the presence of a thin haze layer in the left part of the scene. The maximum difference between the hazy (in the left part of the scene) and clear (the right part) areas is about 5 DNs in TM1. This systematic trend in brightening across the image could be associated with differences in scan sensor viewing of a dense atmospheric moisture/aerosol layer. For comparison, HOT and fourth TC images from TM

P21/R26-b are shown in Figure 7b. There are bright features in the TC transformed image which are obviously not haze-related but rather are associated with high visible-band reflectance ground features such as building complexes, roads and bare ground. On the other hand, the HOT is much less sensitive to these targets.

The image P55/R18 (Figure 7c) covers a mountainous area of mixed forest consisting primarily of mature stands and areas of regeneration following harvesting. The mountain peaks are generally barren rocks, partly covered with ice/snow. A thin stratus cloud layer is present over a broad valley that dominates the middle portion of the scene. The HOT transformation has captured this feature in detail including developing cloud cells along the mountain range embedded in the relatively uniform distributed haze layer. The next three scenes (Figures 7d to 7f) contain complex atmospheric features. In image P49/R21 (Figure 7d), a stratus or cirrostratus layer covers a mountain valley landscape that includes both forest and patches of regenerating low-vegetation. In the sub-image of P34/R23 (Figure 7e), there were two contrails and cumulus cloud clusters over coniferous/deciduous forest and farmland (Figure 7e). In image P15/R29 (Figure 7f), an extensive, thin, transparent cirrus system, in the right part of the image, overlaps bright-white cumulus cloud clusters. In lower part of image P10/R28 (Figure 7g), there are clusters of well-developed thick cumulus clouds with thin ice cloud tops. A thin cirrostratus layer overlaps these convective cloud clusters. In summary, these examples illustrate that diverse atmospheric conditions can be detected in detail by the proposed transformation.

We emphasize again that the HOT is based on the assumption that TM1 and TM3 radiances will be highly correlated for those pixels within the clearest portions of a scene and that this relationship holds for all surface classes. From an inspection of the HOT images for the test TM scenes we conclude that some surface types violate this assumption to varying degrees and could potentially trigger spurious, non-atmospheric, HOT responses. These classes include:

a). Snow cover and shadows over snow. These areas are easily confused with cloudy pixels. Some mountain snow could not be removed from the HOT images (Figures 7c and 7d) due to the fact that in TM1 vs. TM3 spectral space, snow deviates significantly from the linear trend associated with other land spectral classes and lies well into the high haze region of the HOT. Fortunately, this problem can be overcome since snow and cloud can be separated by other means, for example by generating a snow mask based on an analysis of TM band 5 data (Dozier, 1984; Dozier and Marks, 1987).

b). Water bodies. The HOT value over some but not all water surfaces is relatively high. These 'problem' water bodies are apparent in the cases of Figures 7f and 7g. The anomalies might be caused by either sediment in shallow water or by wind-induced waves on the water surface. As with snow, this problem can be easily overcome through the creation of a water body mask, for example, through thresholding the near infrared TM4 image.

c). Bare soil. The response of the HOT over bare soil area appears to be relatively low, as the above modelling results indicate. Under very low haze conditions, fallow fields are dark in the HOT image in comparison to neighbouring crops.

5. HOT Application to Visible Band Adjustment

In this section we describe a methodology to use the HOT image to radiometrically adjust the visible bands of Landsat imagery. The approach is applied the image shown in Figure 7c (P55/R18) to remove the radiometric effects of the broad haze layer that dominates the central portion of scene.

Because of its simplicity, the HOT-based haze image lends itself to a correspondingly direct, first order, methodology for radiometric adjustment. In brief, this procedure involves the following steps.

(a) ***Clear line (CL) definition.*** The scene is visually inspected to locate its clearest regions. Example sub-images in this region and TM1 and TM3 DNs of their constituent pixels are regressed to define the clear line and hence the HOT coefficients. In our example scene, the clear area, labelled C in Figure 11 has been used in the above process.

(b) ***HOT image generation.*** The transformation is applied to each pixel to generate the HOT image (left panel of Figure 7c). The principal haze feature that dominates the central portion of the scene exhibits HOT values in the range 30 to 40 while higher values are associated with surrounding isolated clouds.

(c) ***Histogram generation.*** For each visible band, DN histograms are generated for pixels grouped according to their HOT level. This grouping is in narrow increments of HOT ranging from the clearest to the haziest levels. Figure 8 illustrates some example histograms for band 3 of our example scene. Since, in the visible bands, haze adds to observed radiance, we expect to observe a systematic migration of these histograms to higher DN levels with increasing levels of HOT. We quantify the level of migration of a given histogram by the offset of its lower bound level relative to the histogram lower bound of the clear area pixels. Figure 9 shows the relationship between lower band and HOT for TM bands 1 to 3 where the observed HOT level of the clear area is approximately 30 and histograms have been generated for hazy areas in HOT increments of 1.

(d) ***Radiometric Adjustment Estimation.*** In analogy with methods proposed by others (e.g. Chavez, 1988; Liang et al., 1997), we utilize a form of 'dark target' subtraction to normalize the image to the radiometric level of the clearest areas. To illustrate this procedure, we take an example of TM1. From Figure 9 we note that, for this band, the histogram lower bound for clear pixels (i.e. HOT=30) is approximately 20 DNs. Consider a hazy pixel with an observed HOT level of 40. It is a member of a histogram with a lower bound 27. This implies that the pixel should have its band 1 DN level reduced by 7 during the radiometric adjustment phase. This procedure can be used to adjust all bands for which the histogram analysis has been done. Finally, it should be noted that this form of histogram analysis assumes that similar dark targets are present under all haze conditions, a less stringent condition than full histogram matching required in Richter's (1996b) procedure.

(e) ***Image Adjustment.*** Figure 10 illustrates the results of adjusting TM3 for haze. The lower left image illustrates the whole scene and can be compared with the hazy input image of Figure 7c. Visually the result is pleasing with an apparent good recovery of surface detail under the principal haze feature. In addition, 3 sub-windows have been selected for a more detailed visual 'before-and-after' comparison.

A further quantitative investigation on the effectiveness of the HOT-based method has also been undertaken using the example scene. Four image windows, shown in Figure 11, were selected in the paired hazy and adjusted images for detailed evaluation. One window, labelled 'C' constitutes the clear area used in the CL definition. The other three windows (labelled 1 to 3) are distributed within the dominant haze feature and sample its range of

HOT levels. The four windows include similar land cover types, i.e. mature forests, regenerating harvested areas and associated roads.

The three scattergrams presented in Figure 11 illustrate the effects of adjustment in the TM1-TM3 spectral space. The left plot shows pixels for the clear area (window 'C') only, while the centre and right plots show all pixels within the four windows prior to and following haze adjustment respectively. The impact of the adjustment is to migrate those hazy pixels, constituting the distinct second peak in the center scattergram, back to the clear line. Figure 12 further demonstrates the effectiveness of HOT to characterize haze variation and its impact on TM3. Histograms of HOT values for each window, presented in Figure 12a, illustrate the distinctiveness of the clear vs. hazy regions. The differences in mean HOT value between the hazy windows and the clear window are 7.7, 6.0 and 5.4 (for windows 1 to 3 respectively) while the corresponding differences in mean TM3 DN are 7.8, 6.1 and 4.1 (Figure 12b). By comparing the mean window DN levels before and after adjustment for windows 1 to 3, we estimate that the haze contribution to observed DN level amounted to 15.0, 12.3 and 10.8 per cent for band 3 and 16.9, 13.5 and 11.8 percent for band 1.

6. Conclusions

A robust, Haze Optimized Transform (HOT) has been developed for quantifiying spatial variations in atmospheric contamination on Landsat TM and ETM+ satellite imagery. In a two-dimensional spectral space consisting of visible bands (TM1 vs. TM3), the spectral response of diverse land surface types under clear-sky or low haze conditions are highly correlated and define a distinct 'clear line'. HOT then measures the orthogonal displacement of a pixel from this line. The transform is image-based and easy to use, requiring only that the analyst identify areas within the clearest portions of a scene in order to define the clear line. Its robustness has been validated through its application to seven test TM scenes encompassing a diverse set of atmospheric conditions and surface classes. It should be noted that the methodology is equally applicable to Landsat Multi-Spectral Scanner (MSS) data where the green and red bands can be employed to define a HOT.

Through MODTRAN simulations, the transformation has been investigated and sensitivity studies have been carried out to understand the response of the transform to differing haze levels and haze sources. The main results of this investigation are:

a. The HOT is relatively insensitive to surface cover variations compared to the 4th Tasselled Cap (TC) transform, in particular in reducing the confusion between high haze regions and 'clear-sky' pixels with high visible band reflectance levels.
b. The impact of haze of increasing optical depth can be characterized by a 'trajectory' in spectral space. For differing forms of haze in the TM1-TM3 spectral space these haze trajectories are nearly coincident for the same underlying surface class. This suggests that HOT should be robust and suitable for characterizing a broad range of atmospheric conditions. An application of HOT to a diverse TM image data set confirmed this to be the case.
c. There is monotonic decrease in haze trajectory slope with increasing level of surface reflectance in the visible bands. This phenomenon is of second order and does not preclude the application of an effective first order haze removal procedure that is proportional to the HOT magnitude.

A methodology is presented to adjust visible band radiances for atmospheric contamination based on observed HOT level. A simple, automated procedure can be used in which the relationship between radiometric adjustment level and HOT magnitude is quantified through a histogram analysis procedure analogous to conventional 'dark target' subtraction. The effectiveness of the methodology has been confirmed through its application to one of the test scenes (P55/R18).

Finally, since the compensation procedure results in a migration of pixels back to the clear line in visible spectral space, this has the negative effect of eliminating residual thematic discrimination differences between these bands. In many applications of Landsat data this is an acceptable trade-off. For example, in the creation of the U.S. National Land Cover Data set derived from classification of Landsat TM imagery, only one visible band, TM3, was used (Vogelmann et al. 2001). In the case of Landsat MSS data numerous principal component analyses (e.g. Thomas et al. 1987) have shown that the effective dimensionality of this sensor is really only 2 since the discriminatory power of the sensor arises primarily from differences between visible and infrared spectral responses.

Acknowledgement

The authors wish to thank Dr. Phil Teillet and Mr. Gunar Fedosejevs for providing valuable information and advice on Landsat calibration and atmospheric correction. Dr. Karl Staenz and Dr. Jean-Claude Deguise of the hyperspectral research group at CCRS are also acknowledged for providing spectral signatures used in the simulation studies.

References

Berk, A., Anderson, G.P., Acharya, P.K., Chetwynd, J.H., Bernstein, L.S., Shettle, E.P., Matthew, M.W. and Adler-Golden, S.M. (1999), MODTRAN4 Users Manual. *Air Force Research Laboratory*, pp.94.

Chavez, P.S. (1988), An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24: 450-479.

Chavez, P.S. (1989), Radiometric calibration of Landsat thematic and mapper multispectral images. *Photogrammetric Engineering and Remote Sensing*, 55:1285-1294.

Crist, E.P., and Cicone, R.C. (1984a), A physically-based transformation of thematic mapper data- the TM tasselled cap. *IEEE Transactions on Geoscience and Remote Sensing*. GE22: 256-263.

Crist, E.P. and Cicone, R.C. (1984b), Application of the Tasseled Cap concept to simulated thematic mapper data. *Photogrammetric Engineering and Remote Sensing*. 50: 343-352.

Crist, E.P., Laurin, R., and Cicone, R.C. (1986), Vegetation and soil information contained in transformed thematic mapper data. *Proceedings of the IGARSS' 86 Symposium*, Zurich, Switzerland, 1450-1470.

Dozier, J. (1984), Snow reflectance from Landsat-4 Thematic Mapper. *IEEE Trans. Geosci. Remote Sens*. GE-22: 323-328.

Dozier, J., and Marks, D. (1987), Snow mapping and classification from Landsat Thematic Mapper data. *Ann. Glaciol*. 9: 97-103.

Kaufman, Y.J., (1989), The atmospheric effect on remote sensing and its correction. In *Theory and applications of optical remote sensing* (G. Asrar, Ed), Wiley, New York, p734.

Kauth, R.J., and Thomas, G.S. (1976), The tasselled cap – a graphic description of the spectral-temporal development of the agricultural crops as seen by Landsat. *Proceedings of the Symposium on Machine Processing of Remotely Sensed Data*, Purdue University, 4B41-4B51.

Lavreau, J., (1991), De-hazing Landsat thematic mapper images. *Photogrammetric Engineering & Remote Sensing*. 57: 1297-1302.

Liang, S., Fallah-Adl, H., Kalluri, S., JaJa, J., Kaufman, Y. J., and Townshend, J.R.G. (1997), An operational atmospheric correction algorithm for Landsat Thematic Mapper imagery over the land. *Journal of Geophysical Research.* 102:17173-17186.

Richter, R., (1996a), A spatially adaptive fast \atmospheric correction algorithm. *Int. J. Remote Sensing*. 17: 1201-1214.

Richter, R., (1996b), Atmospheric correction of satellite data with haze removal including a haze/clear transition region. *Computer & Geosciences*. 22: 675-681.

Secker, J., Staenz, K., Gauthier, R.P., and Budkewitsch, P. (2001), Vicarious calibration of airborne hyperspectral sensors in operational environments. *Remote Sensing of Environment*. 76: 81-92.

Song, C., Woodcock, C.E., Seto, K.C., Lenney, M.P., and Macomber, S.A. (2001), Classification and change detection using Landsat TM data: When and how to correct atmosphere effects? *Remote Sensing of Environment*. 75: 230-244.

Teillet, P.M. and Fedosejevs, G. (1995), On the dark target approach to atmospheric correction of remotely sensed data. *Canadian Journal of Remote Sensing*, 21:374-386.

Teillet, P.M., O'Neill, N.T., Kalinauskas, A., Sturgeon, D., and Fedosejevs, G. (1987), A dynamic regression algorithm for incorporating atmospheric models into image correction procedures. *Proceedings of the IGARSS'87 Symposium*, Ann Arbor, Michigan, pp. 913-918.

Thomas, I.L., Benning, V.M. and Ching, N.P. (1987), *Classification of Remotely Sensed Images*, Adam Hilger, Bristol, p. 54.

Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K. and Van Driel, N. (2001), Completion of the 1990s National Land Cover Data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering and Remote Sensing*. 67:650-662.

WMO (1986), A preliminary cloudless standard atmosphere for radiation computation. *WCP-112 WMO/TD*, No. 24: pp53.

Table 1. Surface classes used in the model simulation studies. The surface spectra were retrieved from Probe-1 measurements.

Code	Type	Feature
A	Coniferous forest	Tree canopy
B	Lake water	Median size lake
C	Urban 1	Residential area: mixture of asphalt, concrete, gravel, grass/tree areas, building roof tops, streets and vehicles.
D	Urban 2	Industrial area: mixture of gravel, dirt, buildings, steel bins, trucks, trains, vehicles and vegetation
E	Road	Gravel road intersection with vegetation surrounded
F	Bare soil 1	Dry bare soil
G	Bare soil 2	Wet bare soil
H	Grassland	Grass
I	Deciduous	Deciduous woods and shrubs
J	Cropland 1	Crop in growth
K	Cropland 2	Crop in growth
L	Snow	Unpolluted snow cover

Table 2. Summaries of HOT parameters for the 7 TM test scenes including the clear line (CL) slope (S_{CL}), TM1 and TM3 transform coefficients (SinΘ and CosΘ respectively) and correlation coefficients, r_{CL}.

Scene	Date	Satellite	S_{CL}	SinΘ	CosΘ	r_{CL}
Path21/Row26-a	98-08-12	Landsat-5	1.32	0.797	0.603	0.93
Path21/Row26-b	99-05-27	Landsat-5	1.32	0.797	0.603	0.93
Path55/Row18	99-08-05	Landsat-7	1.57	0.835	0.549	0.95
Path34/Row23	99-08-18	Landsat-7	1.73	0.865	0.500	0.89
Path15/Row29	00-08-15	Landsat-7	1.46	0.825	0.565	0.91
Path10/Row28	99-08-28	Landsat-7	1.72	0.864	0.503	0.87
Path49/Row21	99-09-12	Landsat-7	1.73	0.868	0.496	0.94

Table 3. The types of cloud and surface in TM scenes investigated.

JC: Jet contrail and contrail-induced cloud; Ci: Cirrus; St: thin Stratus or fog; Cu: Cumulus; Mi: mist or thin cirrostratus. A-M of the surface types are corresponding to those in Table 1.

Scene	Cloud type	Surface type
Path21/Row26-a	JC, Ci, Mi	A, B, C, E, H, I
Path55/Row18	St	A, B, I, L
Path34/Row23	JC, Cu	A, B, C, D, E, H, I, J, K
Path15/Row29	Ci, Cu, JC	A, B, C, D, E, F, G, H, I, J, K
Path10/Row28	Cu, Mi, Ci	A, B, C, D, E, F, G, H, I, J, K
Path49/Row21	Ci, Cu	A, B, I, L

FIGURE CAPTIONS

Figure 1. Schematic diagram of the TM1-TM3 spectral space illustrating the conceptual components of the haze optimized transform (HOT). Under clear sky conditions, radiances of common surface cover types, coded as A to H, exhibit high correlation and define a 'clear line' (CL). The effect of haze of increasing optical depth, illustrated by the numerical sequences 1 to 18, is to pixels to 'migrate' away from the clear line. The HOT quantifies the atmospheric contamination level at a pixel location by its perpendicular distance, in spectral space, from the clear line.

Figure 2. Model predictions of HOT (uppercase letters) and 4th Tasselled Cap (TC) transform (lower case letters) values for a mix of surface cover types exhibiting a range in visible reflectances. In both spectral bands, the HOT exhibits a higher degree of insensitivity to surface reflectance, suggesting that it will be more effective in isolating and quantifying atmospheric contamination in images than the TC transform.

Figure 3. Scatterplots of HOT and TC transform values versus TM1 DN level for pixels selected from clear areas of TM scene P21/R26. The model prediction (see Figure 2) of the marked insensitivity of the HOT compared to the TC transform to surface reflectance is confirmed with real imagery.

Figure 4. Scatterplots of HOT and TC transform values versus TM3 DN level for the same pixel sample shown in Figure 3. In this spectral band both transforms exhibit an insensitivity to surface cover.

Figure 5. Comparison of haze 'trajectories' predicted from MODTRAN modelling, for different cloud-only and aerosol-only types. For the five aerosol-only cases, the number is plotted at the data point location. The near coincidence of these trajectories for the same surface cover type suggests that the HOT is robust in responding in a similar way to varying types of atmospheric contamination. On the other hand, it cannot be used to independently estimate optical depth. The surface cover type is coniferous forest.

Figure 6. The simulated radiance for four cloud/aerosol types (water cloud, ice cloud, rural aerosol and urban aerosol) with the same optical depth at 0.55 µm ($\delta_{0.55}$=1.27). The surface cover type is coniferous forest.

Figure 7. Example results of applying HOT to generate a transform image the 7 test scenes. With the exception of Figure 7b, HOT and the parent TM3 images are presented in the left and right panels respectively. In the case of 7b, the 4th Tasselled Cap transform image is displayed in the right panel in order to illustrate the superior performance of HOT in suppressing surface cover response. (7a, 7b, 7c, 7d, 7e, 7f, 7g)

Figure 8. Histograms of TM3 DN level for pixels grouped according to HOT level. The histogram lower bound values are observed to increase with increasing HOT level as expected for the additive effect of haze on image radiometry.

Figure 9. Observed histogram lower bound versus HOT for bands TM1 to TM3 for scene P55/R18. The clear area pixels exhibit a HOT level of approximately 30 while the dominating haze feature of the scene are characterized by HOT values between 30 and 39. The data presented can be used to adjust the radiometry of each band to the clear area level.

Figure 10. Example result of employing the HOT image to radiometrically adjust band 3 of scene P55/R18. Besides the overall scene, selected windows (before and after adjustment images) are presented in full resolution to illustrate recovery of surface detail.

Figure 11. TM1-TM3 scatter plots of four image windows (lower right) with similar surface cover. Window 'C' is located in the clearest portion of the scene while the other three (labelled 1, 2 and 3) are under haze cover. The effect of the adjustment procedure is to bring hazy pixels into alignment with the reference clear area data.

Figure 12. Histograms of (a) HOT and (b) TM3 DN levels for the four windows illustrated in Figure 11. The TM3 data illustrates the data migration arising from the adjustment process.