ELSEVIER

Radiometric cross-calibration of the Landsat-7 ETM+ and Landsat-5 TM sensors based on tandem data sets

P.M. Teillet[a,*], J.L. Barker[b], B.L. Markham[b], R.R Irish[c], G. Fedosejevs[a], J.C. Storey[d]

[a]Canada Centre for Remote Sensing, 588 Booth Street, Ottawa, Ontario, Canada K1A 0Y7
[b]NASA's Goddard Space Flight Center, Code 923, Greenbelt, MD 20771, USA
[c]Science Systems and Applications, Inc., 5900 Princess Garden Parkway, Lanham, MD 20706, USA
[d]Raytheon Information Technology and Systems Services, 4400 Forbes Boulevard, Lanham, MD 20706-4392, USA

Abstract

Early in its mission, the Landsat-7 spacecraft was temporarily placed in a "tandem" orbit very close to that of the Landsat-5 spacecraft in order to facilitate the establishment of sensor calibration continuity between the Landsat-7 Enhanced Thematic Mapper Plus (ETM+) and Landsat-5 Thematic Mapper (TM) sensors. The key period for the tandem configuration was June 1–4, 1999, during which hundreds of nearly coincident matching scenes were recorded by both the Landsat-7 ETM+ and, in cooperation with Space Imaging/EOSAT and international ground stations, the Landsat-5 TM as well. The paper presents a methodology for radiometric cross-calibration of the solar reflective spectral bands of the Landsat-7 ETM+ and Landsat-5 TM sensors and results based on analysis of two different tandem image pairs for which ground reference data are available. With the well-calibrated ETM+ as a reference, the tandem-based cross-calibrations for the two image pairs yield TM responsivities that are consistent to each other to within 2% or better depending on the spectral band. Comparisons with independent methods and results obtained by other groups indicate that the tandem-based cross-calibration is within 3% of the independent results on average in spectral bands 1–4 but compares less favorably in bands 5 and 7. The present study indicates that the tandem cross-calibration approach can provide a valuable "contemporary" calibration update for Landsat-5 TM in the visible and near-infrared spectral bands based on the excellent radiometric performance of Landsat-7 ETM+. The methodology also incorporates adjustments for spectral band differences between the two Landsat sensors. Spectral band difference effects are shown to be more dependent on the surface reflectance spectrum than on atmospheric and illumination conditions. A variety of terrestrial surfaces are assessed regarding their suitability for Landsat radiometric cross-calibration in the absence of surface reflectance spectra. Crown Copyright © 2001 Elsevier Inc. All rights reserved.

1. Introduction

1.1. Calibrating the Landsat data record

Earth observation imagery from the Landsat series of sensors spans almost three decades beginning in 1972. The image data acquired by the Multispectral Scanners (MSS) and the Thematic Mappers (TM) represent a unique historical record of the Earth's surface at spatial scales of 30 to 80 m. MSS sensors were carried on-board Landsats 1 through 5 (1972–present) and TM sensors were carried on-board Landsat-4 (1982–1993) and Landsat-5 (1984–present).

MSS and TM data coverage acquired by the US has been global with the exception of the years following the loss of access to the Tracking and Data Relay Satellite System due to the Ku-band transmitter failure. Some of this gap has been filled through data reception by international ground stations. Global data acquisition by the US has resumed with the 1999 launch of the Landsat-7 Enhanced Thematic Mapper Plus (ETM+) and its solid-state recorders.

The Landsat data record is important for terrestrial remote sensing and global change research because of its relatively fine spatial resolution, extensive terrestrial coverage, and temporal baseline over a time when significant anthropogenic terrestrial change has occurred. In order to benefit fully from such a data record, steps are needed to ensure that the data are self-consistent and not significantly affected by artefacts of the measurement system. A critical step in such a process is sensor radiometric calibration to an

---

absolute scale, yielding image data at the top of the atmosphere in physical units. Additional processing steps to retrieve surface parameters such as reflectance and temperature then become possible.

Considerable effort went into the design and development of the Landsat sensors, resulting in some of the most stable Earth observation sensors ever. The 17-year performance of the Landsat-5 TM is a testimony to the durability and quality of the sensor. Prelaunch characterizations were extensive and the Landsat Image Data Quality Assessment program in the early postlaunch years of the Landsat-4/5 sensors was successful in achieving on-orbit characterizations. With the success of Landsat-7, renewed efforts are underway to ensure radiometric calibration across the Landsat series of sensors and with other Earth observation sensors.

1.2. Radiometric cross-calibration

Approaches to sensor radiometric calibration and cross-calibration have been well-documented (Dinguirard & Slater, 1999) and new methodologies continue to evolve (Teillet et al., 2001). Briefly, consistency between different sensors starts with sound calibration of the individual sensors, including the development of a stable sensor, detailed prelaunch characterization, and on-orbit calibration. Postlaunch radiometric calibrations can be based on reference to onboard standards, solar and lunar illumination, and ground-based test sites. Cross-calibration between sensors can be based on prelaunch measurements in the laboratory using common sources or transfer radiometers at the same or different times. For those missions operating during the same time periods, postlaunch cross-calibration can make use of near-simultaneous imaging of common targets on the surface of the Earth or Moon or mutual reference to pseudo-invariant features or data from a third sensor. Typically, for any given series of satellite sensor systems, there is a limited overlap period when more than one of the sensors is operating at the same time. Such an overlap period was designed into the initial phases of the Landsat-7 mission. The resulting opportunity for radiometric cross-calibration is the main subject of this paper.

1.3. The tandem configuration

The launch of Landsat-7 on April 15, 1999, placed the spacecraft temporarily in an orbit very close to that of the Landsat-5 spacecraft. The mean altitude of Landsat-7 was 699 km, 6 km below the 705-km mean altitude of Landsat-5. At this altitude, the Landsat-7 ground track drifted slowly relative to the nearly fixed Landsat-5 pattern. The key period for the tandem configuration was June 1–4, 1999, when the tracks were almost exactly the same, but with a temporal offset on the order of 10 to 30 min. This unusual and valuable opportunity was specifically designed to facilitate the establishment of data consistency between the Landsat ETM+ and TM sensors. During the tandem configuration period when there was useful overlap in coverage between the two sensors, image sequences corresponding to 791 matching scenes were recorded by both the Landsat-7 ETM+ and, in cooperation with Space Imaging EOSAT and international ground stations, the Landsat-5 TM (Fig. 1 and Table 1). Subsequently, the Landsat-7 orbit was changed for nominal operations such that its 16-day repeat coverage cycle is now offset from that of Landsat-5 by 8 days. Given cloud cover and possible problems with data reception and recording, the number of useful tandem data scene pairs is roughly estimated to be on the order of 400 scenes.

This paper describes a cross-calibration methodology applicable to tandem image pairs and presents specific results for two different pairs of nearly coincident matching scenes from the tandem configuration period. The main results consist of TM responsivities in the six solar reflective spectral bands referenced against well-calibrated ETM+ responsivities in corresponding spectral bands. The formulation includes adjustments for differences in illumination regimes as well as for differences in spectral response profiles between the two sensors.

2. Tandem data sets selected for analysis

Attention was focused on two particular tandem image pairs for cross-calibration methodology development and analysis because of the availability of ground reference data. Both Landsat sensors imaged the Railroad Valley Playa, Nevada (RVPN) on 1 June 1999, when a team from the University of Arizona (UAZ) made measurements of surface spectral reflectance and atmospheric aerosol optical depth the same day. Similarly, a team from South Dakota State University (SDSU) acquired the same types of ground reference data at a grassland test site in the area of Niobrara, Nebraska (NIOB) on 2 June 1999, the day of the tandem Landsat overpasses for that site.

Table 2 provides information on the characteristics of the two data sets and Fig. 2 shows both Landsat image pairs. The RVPN test site is a dry-lake playa that is very homogeneous and consists of compacted clay-rich lacustrine deposits forming a relatively smooth surface compared to most land covers. The NIOB test site is characterized primarily by grasslands grazed by cattle and by a smaller proportion of agricultural crops.

3. Cross-calibration methodology based on tandem data sets

The following methodology takes advantage of the opportunity presented by tandem data sets to cross-calibrate the solar-reflective spectral bands of the Landsat-5 TM and Landsat-7 ETM+ sensors. A similar opportunity arose after
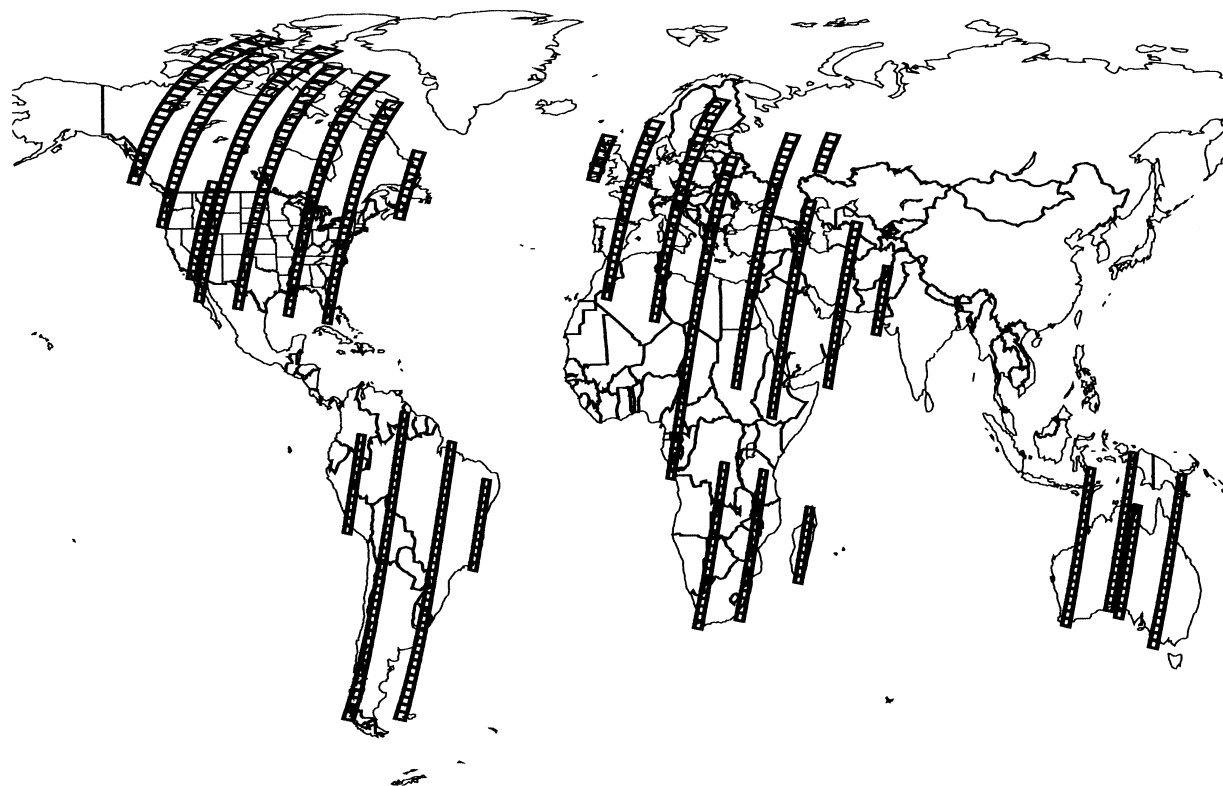
Fig. 1. Landsat-7 ETM+ and Landsat-5 TM coincident acquisitions during the tandem configuration period (cylindrical equidistant projection).

Table 1
Landsat-7 ETM+ and Landsat-5 TM tandem data coverage (see also Fig. 1)

Tandem scene coverage (June 1–4, 1999)

June	Path	Row	Station	June	Path	Row	Station
3	6	21–29	GNC	3	159		RSA
3	6	57–71	CUB	3	159	69–78	JSA
3	6	67–71	COA	2	168	19–27	KIS
2	15	12–44	GNC	2	168		RSA
2	15	27–45	NOK	2	168	62–83	JSA
3	22	10–43	GNC	3	175	19–26	KIS
3	22	26–49	NOK	3	175	23–42	FUI
2	31	7–40	PAC	3	175		RSA
2	31	25–46	NOK	3	175	62–85	JSA
3	38	6–39	PAC	2	184	22–44	FUI
3	38	25–40	NOK	2	184	44–77	LBG
1	40	25–38	NOK	3	191	14–24	KIS
2	47	4–30	PAC	3	191	17–43	FUI
2	47	25–30	NOK	2	200	17–40	FUI
3	54	4–25	PAC	3	207	19–24	FUI
2–3	95	65–87	ASA	2	216	63–76	CUB
3	102	69–83	ASA	3	223	60–86	CUB
2	104	62–82	ASA	3	223	68–98	COA
3–4	111	64–84	ASA	2	232	54–85	CUB
2	152		RSA	2	232	66–97	COA

Stations: ASA = ACRES, Alice Springs, Australia; COA = Cordoba, Argentina; CUB = INPE, Cuiaba, Brazil; FUI = ESA, Fucino, Italy; GNC = CCRS, Gatineau, Canada; JSA = Johannesburg, South Africa; KIS = ESA, Kiruna, Sweden; LBG = DLR, Libreville, Gabon; NOK = SI/EOSAT, Norman, OK, USA; PAC = CCRS, Prince Albert, Canada; RSA = Saudia Arabia (for SI/Dubai).

the launch of Landsat-5 when its orbit was temporarily close to that of Landsat-4. Metzler and Malila (1985) cross-calibrated one tandem pair of fully processed scenes, i.e., scenes that had undergone radiometric and geometric correction as opposed to the raw data used in the present work.

Table 2
Characteristics of the two data sets, where AOD550 represents aerosol optical depth at 0.550 μm.

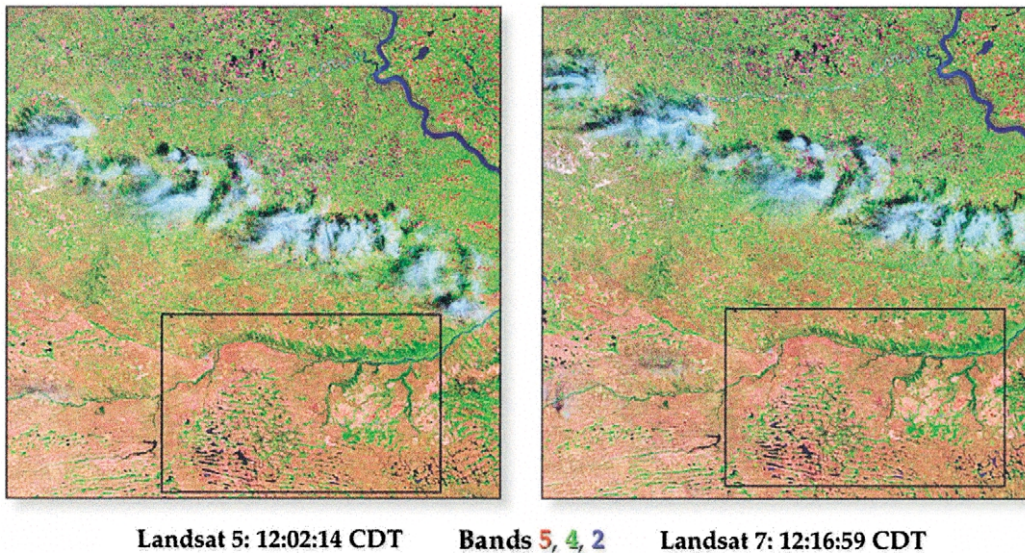|	RVPN	NIOB
Image date	01 June 1999	02 June 1999
WRS path/row	40/33	31/30
Landsat-7 offset from WRS	76.56 km east	18.15 km east
ETM+ data level	Level-0R	Level-0R
TM data level	Level-0	Level-0
ETM+ solar zenith angle	24.28°	26.60°
TM solar zenith angle	27.23°	28.67°
Terrain elevation	1.425 km	0.760 km
ETM+ AOD550	0.1046	0.059
TM AOD550	0.1035	0.059
Area common to ETM+ and TM	10.7 × 4.4 km	106 × 66 km
ETM+ grid cell size	71 pixels by 29 lines	696 pixels by 434 lines
ETM+ pixels/cell	2059	308,352
TM grid cell size	71 pixels by 29 lines	704 pixels by 438 lines
TM pixels/cell	2059	302,064

Railroad Valley Playa, Nevada, WRS 40/33, June 1, 1999

Landsat 5   Bands 5, 4, 2   Landsat 7
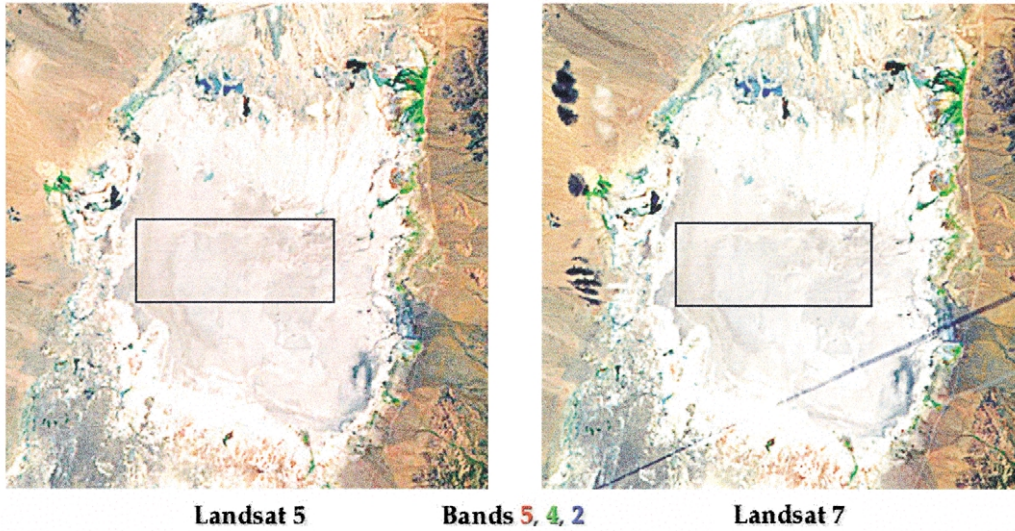
Niobrara, Nebraska, WRS 31/30, June 2, 1999

Landsat 5: 12:02:14 CDT   Bands 5, 4, 2   Landsat 7: 12:16:59 CDT

Fig. 2. Two different tandem image pairs used for cross-calibration between Landsat-5 TM and Landsat-7 ETM+. The RVPN subscenes are 20 km across whereas the NIOB subscenes are 120 km across.

Price (1989) utilized a simpler methodology compared to that of Metzler and Malila, but the two methods gave similar results. Spectral band difference effects were not an issue in the cross-calibrations between the TM sensors on Landsat-4 and Landsat-5, whereas they are a critical aspect of the Landsat-7 and Landsat-5 cross-calibration.

The present approach assumes that the Landsat-5 TM calibration is to be updated with respect to the Landsat-7 ETM+ sensor, which serves as a well-calibrated reference sensor with a radiometric calibration uncertainty of ±3% (Barker et al., 2000). Because data acquisitions were only 10 to 30 min apart during the tandem configuration period,

it is assumed that the surface and atmospheric conditions did not change significantly between the two image acquisitions. Nevertheless, there are geometric, radiometric, and spectral considerations to be addressed.

3.1. Geometric matching

Geometrically, the Landsat-7 and Landsat-5 sensors differ in their along-track and across-track pixel sampling. Due to wearing of the bumpers used by the Landsat-5 TM scanning mirror, along-track gaps between scans are longer than they are for Landsat-7 ETM+. For the same reason and

because the ETM+ scan time is slightly longer than the specification, there are also across-track differences in the ground coverage. In addition, slight mismatches will arise in the imagery because of the altitude difference. In particular, there is variation in the ETM+ scanning pattern and its effect on the scan line corrector due to the lower-than-nominal orbit during the tandem configuration time period. These considerations make it very difficult to establish sufficient geometric control to facilitate radiometric comparisons on a point-by-point and/or detector-by-detector basis. Therefore, the analysis approach was developed to make use of image statistics based on large areas in common between the image pairs.

3.2. Radiometric formulation

Radiometrically raw data are assumed (Level 0 for TM and Level 0R for ETM+). In spectral band i, the image quantized level Q_i (in counts) is related to top-of-atmosphere (TOA) radiance L_i^* [in W/(m^2 sr μm)] by Eq. (1)

$$Q_i = G_i L_i^* + Q_{0i}, \tag{1}$$

where G_i is band-averaged sensor responsivity (in counts per unit radiance, CPUR) and Q_{0i} is the zero-radiance bias (in counts) in spectral band i. Quantized levels of Q = 0 and Q = 255 are excluded to avoid saturation effects. The zero-radiance biases are based on DC restore values computed on a line-by-line basis. Radiometric detector normalizations based on full-scene statistics are applied in each spectral band, for each particular scene in the case of TM and for many scenes in the case of ETM+. The normalizations are with respect to the band average and the process is not expected to bias the cross-calibration. Normalized and bias-corrected image values are then given by

$$\Delta Q_i = Q_i - Q_{0i} = G_i L_i^*. \tag{2}$$

Thus, TOA radiances L_i^* [in W/(m^2 sr μm)] are related to image data by

$$L_i^* = \Delta Q_i / G_i. \tag{3}$$

TOA reflectance is related to TOA radiance by

$$\rho_i^* = \pi L_i^* d_s^2 / (E_{0i} \cos\theta), \tag{4}$$

where E_{0i} is the exo-atmospheric solar irradiance in spectral band i [in W/(m^2 μm)] based on the Modtran-3 spectrum, θ is the solar zenith angle, and d_s is the Earth–Sun distance in astronomical units. A combination of Eqs. (2)–(4) yields

$$\Delta Q_i = G_i \rho_i^* E_{0i} \cos\theta / (\pi d_s^2). \tag{5}$$

There are two advantages to using reflectances instead of radiances. One advantage is to remove the cosine effect of different solar zenith angles due to the 10- to 30-min time difference between data acquisitions. For example, the 3° difference in solar zenith angles for the RVPN image pair leads to a 2.5% effect in the ratio of the cosines of the

respective angles. The other advantage is to compensate for different values of exo-atmospheric solar irradiance arising from spectral band differences. If differences in atmospheric conditions are not a factor, then the TOA reflectance comparisons have the potential to yield the best possible calibration comparisons between the TM and ETM+ based on the tandem data sets.

Several sensor-related considerations have not been addressed in this study, although their neglect is not expected to be a significant source of error in the TM radiometric calibration update. Detector-specific, scan-direction, and scene-to-scene within-orbit behaviors have not been considered. Comparisons have not been made between ETM+ data from the tandem configuration period versus ETM+ data acquired and processed after the change to routine operations.

Cross-calibration methodologies in general should consider adjustments as appropriate for bidirectional reflectance factor (BRF) effects due to differences in illumination and observation angles. Even if the same test sites are imaged the same day, significant overpass time and off-nadir viewing differences can arise depending on the satellite sensor. For Landsat sensor image data pairs acquired during the tandem configuration period, the expectation is that such BRF adjustments are not necessary. The solar illumination geometries are very similar (within 3°), satellite zenith angles are predominantly near-nadir, and relative azimuth angles between solar and satellite directions do not differ significantly from one Landsat overpass to the other.

3.3. Cross-calibration

Eq. (5) can be defined separately for image data from the Landsat-5 TM ("5") and for image data from the Landsat-7 ETM+ ("7"):

$$\Delta Q_{i5} = G_{i5} \rho_{i5}^* (E_{0i} \cos\theta)_5 / (\pi d_s^2); \tag{6}$$

$$\Delta Q_{i7} = G_{i7} \rho_{i7}^* (E_{0i} \cos\theta)_7 / (\pi d_s^2). \tag{7}$$

The combination of Eqs. (6) and (7) yields

$$\Delta Q_{i5A} = A_i \Delta Q_{i5} = (G_{i5}/G_{i7}) \Delta Q_{i7} = M_i \Delta Q_{i7}, \tag{8}$$

where the adjustment factor A_i adjusts Landsat-5 TM radiance data for illumination and spectral band difference effects (Table 3). In particular, M_i is the slope of the linear equation that characterizes ΔQ_{i5A} as a function of ΔQ_{i7} and

$$A_i = B_i (E_{0i} \cos\theta)_7 / (E_{0i} \cos\theta)_5, \tag{9}$$

where

$$B_i = \rho_{i7}^* / \rho_{i5}^*. \tag{10}$$

B_i is essentially a spectral band adjustment factor, given that the ρ_i^* in Eqs. (6) and (7) are not necessarily the same because of the differences in relative spectral response

Table 3
Spectral band quantities used in the cross-calibration analysis

Spectral band	1	2	3	4	5	7
TM E_o [W/(m^2 μm)]	1954	1826	1558	1047	217.2	80.29
ETM+ E_o [W/(m^2 μm)]	1968	1839	1555	1054	228.4	81.59
A_i for RVPN	1.013	1.013	1.017	1.035	1.106	0.994
B_i for RVPN	0.981	0.981	0.994	1.003	1.026	0.954
A_i for NIOB	1.051	1.053	0.997	1.044	1.118	0.969
B_i for NIOB	1.020	1.022	0.977	1.014	1.039	0.932

E_o is the exo-atmospheric solar irradiance, A_i is the illumination adjustment factor, and B_i is the spectral adjustment factor.

profiles between corresponding ETM+ and TM spectral bands (discussed in the next section). Landsat-5 TM responsivity G_{i5} is then given in spectral band i (in CPUR) by

$$G_{i5} = M_i G_{i7}. \tag{11}$$

With this updated value of TM responsivity, users can obtain TOA radiance L_i^* [in W/(m^2 sr μm)] from raw image quantized levels Q_i (in counts) using

$$L_i^* = a_i Q_i + b_i, \tag{12}$$

where $a_i = 1/G_i$ and $b_i = -Q_{0i}/G_i$. Thus, image pairs from the tandem configuration period make it possible to use well-calibrated Landsat-7 ETM+ image data to update the radiometric calibration of the Landsat-5 TM.

3.4. Spectral band differences

There are significant differences in relative spectral response profiles between corresponding Landsat-7 ETM+
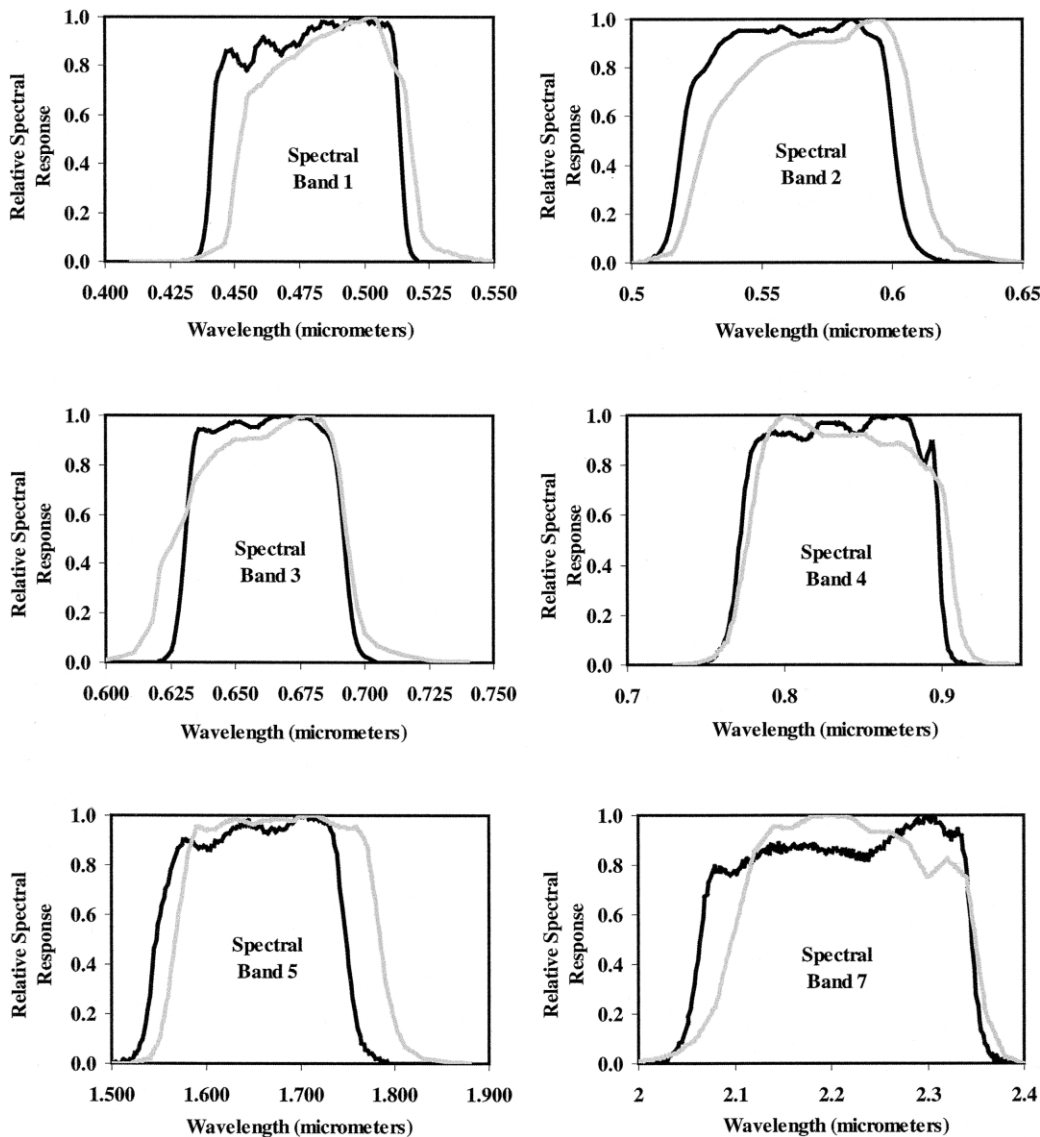

Fig. 3. Relative spectral response profiles that show the spectral band differences between Landsat-7 ETM+ and Landsat-5 TM. The ETM+ and TM profiles are the black and grey curves, respectively.

Railroad Valley Playa, June 1, 1999 (UA)

Niobrara, Nebraska, June 2, 1999 (SDSU)

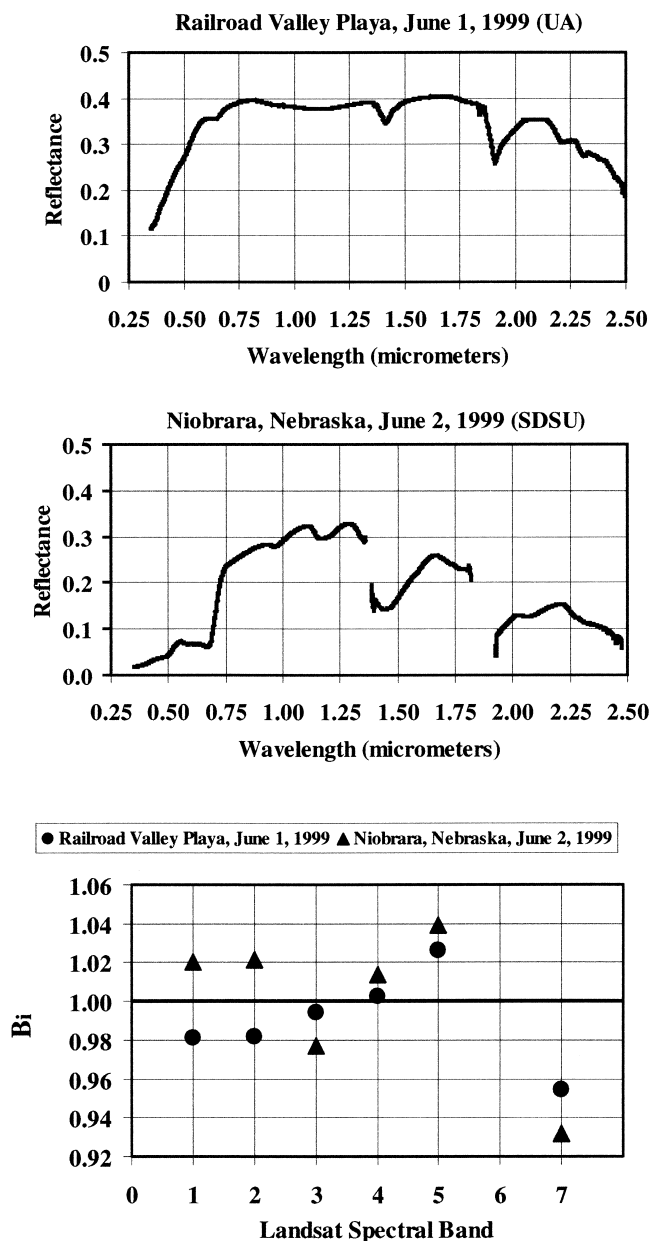● Railroad Valley Playa, June 1, 1999 ▲ Niobrara, Nebraska, June 2, 1999

Fig. 4. The top two plots are surface reflectance spectra for the RVPN and NIOB test sites as acquired by the UAZ and SDSU, respectively. The bottom plot shows the spectral band adjustment factors, B_i, for the two test sites as a function of spectral band i.

and Landsat-4/5 TM spectral bands (Fig. 3). The effects these spectral band differences have on measured TOA reflectances depend on spectral variations in the exo-atmospheric solar illumination, the atmospheric transmittance, and the surface reflectance. Because surface spectral reflectance and atmospheric aerosol optical depth data are available for the RVPN and NIOB test sites on June 1 and 2, 1999, respectively, B_i factors were generated for the image data pairs under consideration in this study (Fig. 4). For each test site, these data were used as inputs to a radiative transfer code to compute the TOA reflectance in corresponding solar reflective ETM+ and TM band. Standard values were

assumed for other atmospheric input parameters (mid-latitude summer profile and continental aerosol model), but solar illumination angles pertinent to each Landsat image acquisition under consideration were used (Table 2). The results in Fig. 4 show that the spectral band difference effect is on the order of 2%, except in the two shortwave infrared bands where it is larger, ranging from 3% to 7% depending on the band and the test site. The direction of the effect is also opposite between the two sites in bands 1 and 2, which is attributable to the significantly different reflectance spectra of the playa and grassland surfaces. Deriving spectral adjustment factors for other tandem data pairs would be less straightforward because of the lack of ground reference data.

In order to assess the magnitude of the spectral band difference effect for a variety of surface cover types, a series of simulations were carried out using the CAM5S atmospheric radiative transfer code (O'Neill, Royer, & Nguyen, 1996). Reflectance spectra for six surface cover or target types were used: bright vegetation, black spruce, rangeland, dry sand, clear water, and 2-day-old snow. Other conditions and parameters involved in the simulations are given in Table 4. The main input parameters are two levels of aerosol optical depth at 0.550 μm (AOD550 = 0.05 for a very clear atmosphere and 0.5 for a hazy atmosphere with a visibility of 6 km), two different solar zenith angles (SZA = 15° and 60°), and two standard atmospheric models [dry (US62) and humid (tropical)]. The code outputs were TOA reflectances in the solar reflective spectral bands for both the ETM+ and TM. B_i spectral adjustment factors were then computed (Eq. (6)) and the various simulation results compared.

Black spruce comparisons are shown as examples in Fig. 5. The comparisons for the other target types (not shown) yield similar or even smaller differences in spectral band difference effects. The results for all target types examined clearly indicate that atmospheric and illumination conditions generally contribute significantly less to the

Table 4
Conditions and parameters involved in TOA reflectance simulations carried out to assess spectral band difference effects

Landsat sensor	Landsat-7 ETM+; Landsat-5 TM
Spectral bands	1, 2, 3, 4, 5, 7
Aerosol optical depth at 0.550 μm (AOD550)	0.05; 0.5
Atmospheric model	US62: U(H₂O) = 1.424 g/cm²; U(O₃) = 0.344 cm-atm Tropical: U(H₂O) = 4.12 g/cm²; U(O₃) = 0.247 cm-atm
Solar zenith angle (SZA)	15°; 60°
Surface reflectance spectrum	Bright vegetation, black spruce, rangeland, dry sand, clear water, 2-day-old snow
Conditions common to all cases	Sea-level terrain elevation, nadir viewing geometry, Earth–Sun distance = 1 AU, continental aerosol model

U(H₂O) = water vapor content. U(O₃) = ozone content. AU = astronomical units.
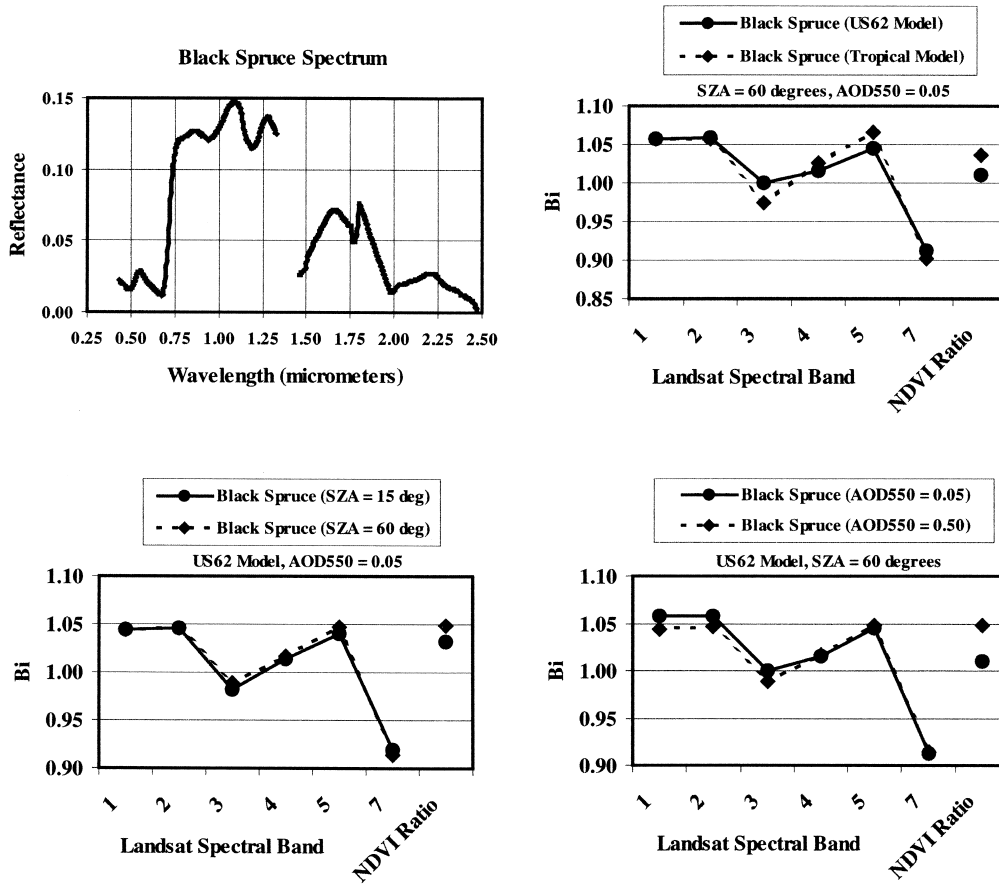
Fig. 5. Comparisons of spectral band adjustment factors, B_i, and NDVI ratio for the indicated atmospheric and illumination conditions for the black spruce spectrum (shown in the upper-left plot). The upper-right plot compares results for dry (US62) and humid (tropical) atmospheric models. The lower-left plot compares results for solar zenith angles of 15° and 60°. The lower-right plot compares results for aerosol optical depths of 0.05 and 0.5 at 0.550 μm. Points are connected by lines for visualization purposes only.

spectral band difference effect than does the surface reflectance spectrum itself. The low sensitivity to atmospheric and illumination conditions is consistent with expectations. Aerosol scattering varies slowly with wavelength and lower aerosol optical depths allow the stronger wavelength dependence of molecular (Rayleigh) scattering to increase slightly the spectral band difference effect. Landsat spectral bands are well placed to avoid most atmospheric absorption features and hence there is not a lot of difference in general between results for dry and humid atmospheric conditions. Nevertheless, the effect of differences in the water vapor content approaches 2% in a few spectral bands.

Fig. 6 shows the results for all target types involved in the simulations plotted in two groups, in all cases for SZA = 60°, AOD550 = 0.05, and a US62 atmospheric model. The RVPN and NIOB grassland cases are also included and based on their own specific atmospheric and illumination conditions as measured during the university field campaigns. B_i factors for bright vegetation, black spruce, and clear water (Fig. 6, top) indicate significant spectral band difference effects, reaching as high as 10% and averaging around 5%. The implication is that dense vegetation cover and water should not be used for radio-

metric cross-calibration unless surface reflectance spectra are available. The spectral band difference effects for the other five target types (Fig. 6, bottom) are all within 2.5% for spectral bands 1–4. Thus, rangeland, grassland, sand, playa, and snow are potentially good candidates for radiometric cross-calibration in the visible and near-infrared bands even in the absence of surface reflectance spectra. However, cross-calibration in the two shortwave infrared bands remains a problem without surface reflectance spectra. Although the rangeland and grassland spectra used in this study show minimal spectral band difference effects, they are still vegetated surfaces that can undergo significant phenological change. Therefore, caution is advised in the use of such target types for radiometric cross-calibration without ground reference data.

Fig. 6 also includes the ratio of the Normalized Difference Vegetation Index (NDVI) obtained for ETM+ and TM (Eq. (13)), where

$$NDVI = (\rho_4^* - \rho_3^*)/(\rho_4^* + \rho_3^*). \tag{13}$$

Due to the spectral band difference effect, NDVI results are higher for the ETM+ for all four target cases considered,
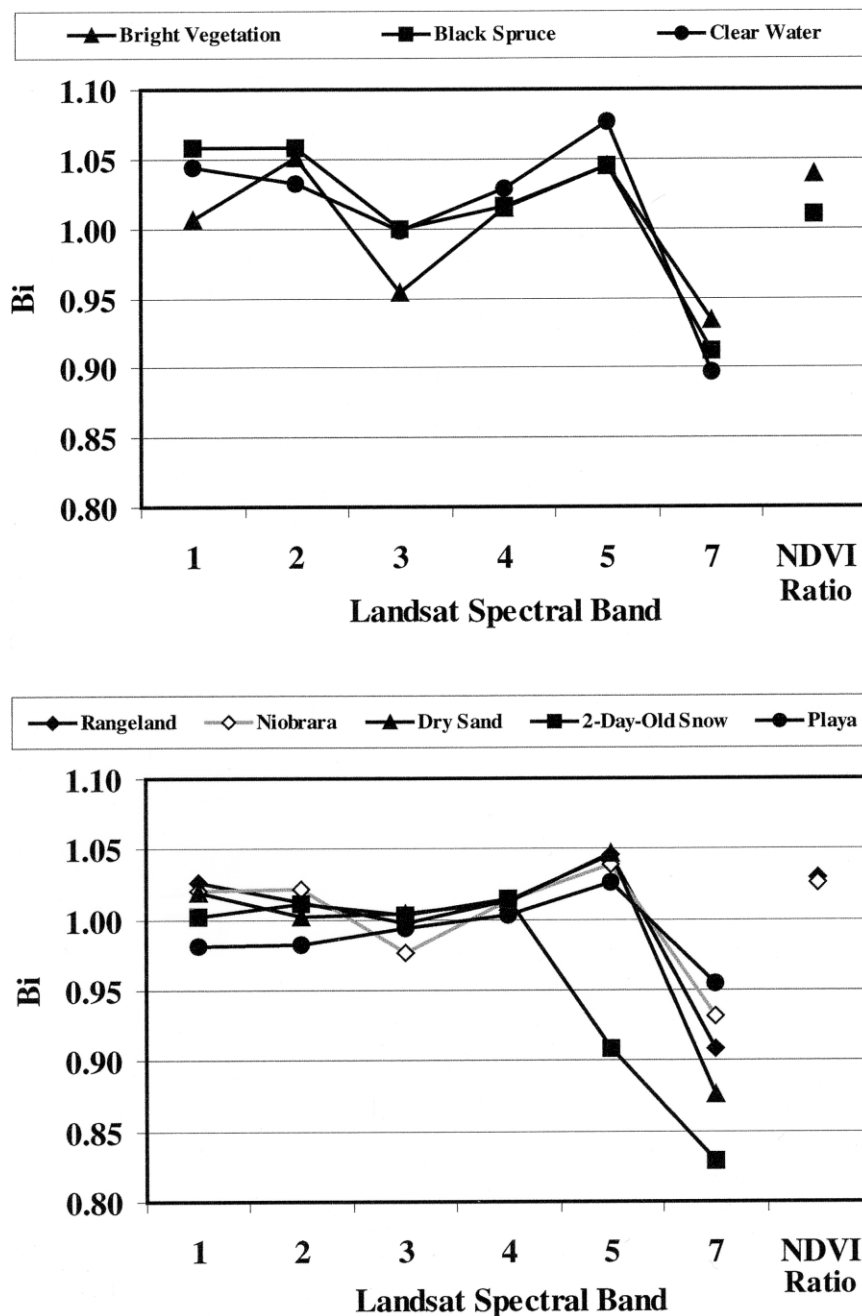
Fig. 6. Comparisons of spectral band adjustment factors, B_i, and NDVI ratio for various target types: (top) bright vegetation, black spruce, and clear water; (bottom) rangeland, NIOB grassland, dry sand, 2-day-old snow, and RVPN. Points are connected by lines for visualization purposes only.

ranging from 1% to 4% higher and averaging about 2.5% higher.

4. Image processing and analysis

Standard image processing and statistical analysis steps were used to obtain the M_i slopes in Eq. (8) for use in Eq. (11). The ΔQ_{i5A} and ΔQ_{i7} for use in Eq. (8) were obtained from large areas in common between ETM+ and TM data pairs. As noted before, subpixel geometric registration is not

critical in this case, but care was taken to capture the common area as accurately as possible.

The image processing steps in each solar-reflective spectral band i were as follows.

1. Set up a 5×5 grid of contiguous image windows or cells and extract ΔQ_i means and standard deviations from each of the 25 grid cells for an area common to both the ETM+ and TM image data (Fig. 7).
2. Repeat Step 1 for a series of 1-pixel shifts in a 5×5 pattern, yielding 25 subsets of means and standard

71 pixels by 29 lines (2.1 km by 0.9 km)

Landsat-7 ETM+ Sub-Scene Spectral Band 1

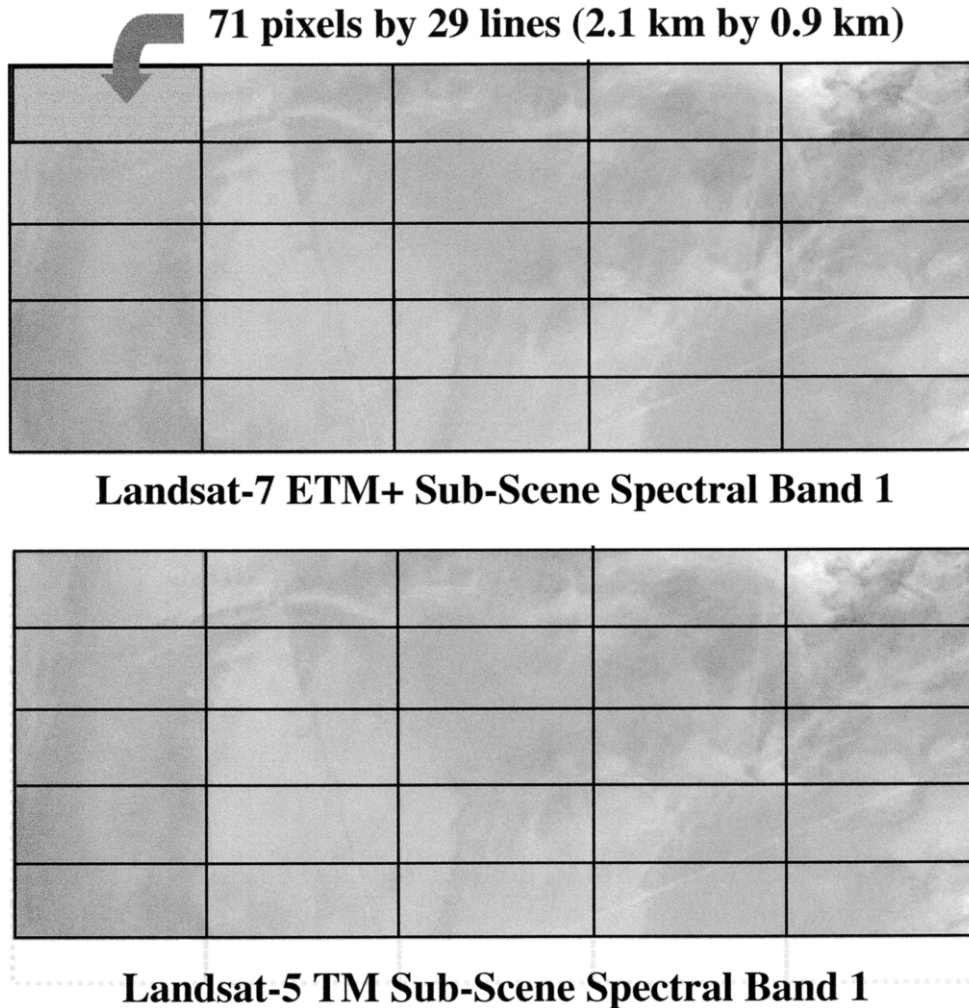**Landsat-5 TM Sub-Scene Spectral Band 1**

Fig. 7. Grid cell analysis scheme illustrated for the RVPN test site, 1 June 1999.

deviations per grid cell. This "jitter" pattern makes it possible to assess the sensitivity of grid cell data to these shifts as an indicator of misregistration effects. The additional sets of values resulting from this jitter exercise are not used for any other purpose.

3. Keep grid cell mean results only if sensitivity to shifts is low (within 1%). The value retained is the one obtained for the geometric center of the jitter pattern.
4. Compute ΔQ_{i5A} from ΔQ_{i5} using Eqs. (8)–(10) to adjust for spectral band differences and illumination regime differences between acquisitions.
5. Plot grid cell ΔQ_{i5A} and ΔQ_{i7} means and obtain the slopes M_i (Eq. (8)).
6. Use Eq. (11) to compute Landsat-5 TM responsivity G_{i5}.

Table 2 summarizes the common areas and grid cell sizes for each test site image pair. For the NIOB subscene, the grid covered as large a common area of grassland as possible excluding clouds and cloud shadows. Scene portions down-scan from bright targets such as clouds were excluded in order to reduce the impact of the TM memory effect. It should be noted that an area common to the two images in a pair can have slightly different numbers of lines and pixels because of differences in satellite altitude and sensor scanning times. For the NIOB image, the overall grid of 5×5 cells covers approximately 106×66 km. Specifically, each ETM+ grid cell encompasses 696 pixels by 434 lines and each TM grid cell encompasses 704 pixels by 438 lines. In either case, there are over 300,000 pixels per grid cell. For the RVPN subscene, a smaller common area was selected since the playa occupies a limited portion of the image. In that case, the overall grid was selected to be 10.7×4.4 km, covering the central part of the playa, such that the grid cell size is 71 pixels by 29 lines (2059 pixels per cell).

The jitter exercise revealed low sensitivity to possible misregistration between the subscenes selected as areas common to both images in each tandem pair. For the majority of grid cells, the coefficient of variation for the 25 jitter values of ΔQ_i is a small fraction of a percent, reaching 0.3% in a few cases for the NIOB scene and just
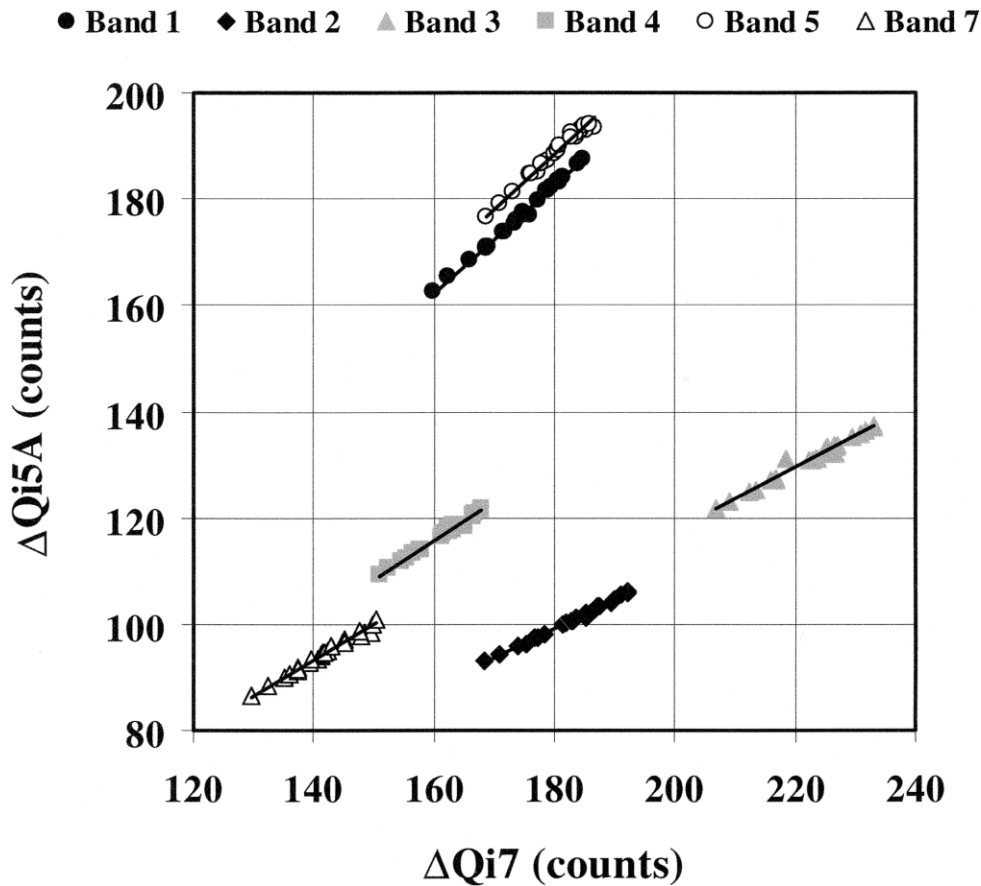
Fig. 8. Plot of grid cell ΔQ_{i5A} and ΔQ_{i7} means for the RVPN subscenes. The lines are linear fits with zero-intercepts.

over 1% for one grid cell for the RVPN scene. The average coefficient of variation for the RVPN case is 0.24%. Therefore, no image cells were excluded on the basis of the jitter exercise.

After preprocessing in accordance with the radiometric formulation described earlier, subscene grid cell means for ΔQ_{i5A} and ΔQ_{i7} were plotted to obtain the slopes M_i (Eq. (8)). Figs. 8 and 9 show plots for RVPN and Table 5 lists the slope results for the RVPN and NIOB subscene pairs analyzed separately and in combination (Fig. 10). Because the quantized levels are bias-subtracted, the linear fits were forced to have zero intercepts. Linear fits were also obtained without this constraint and intercepts of several digital counts were found in some spectral bands (Table 5). Nevertheless, with the exception of band 4 for the NIOB case, the unaccounted for variances in percent, $100 (1 - R^2)$, with the linear fits are low (Table 5), where R is the correlation coefficient. No explanation has been found for the greater scatter in spectral band 4. Table 5 also indicates that the M_i slopes obtained for the two different image pairs generally differ by a few percent only, which provides some degree of confidence in the cross-calibration methodology.

Fig. 9 presents the RVPN results from Fig. 8 in terms of residuals from linear fits with zero intercepts. Almost all residuals are well within the $\pm 1\%$ range and exhibit no

systematic trends. Residuals for the NIOB case (not shown) reveal no systematic trends. The residuals are almost all within $\pm 1\%$ for spectral bands 1, 2, 5, and 7. Band 3 residuals are almost all within $\pm 1.5\%$. There is more scatter in the band 4 residuals, averaging approximately $\pm 2.5\%$, but with several points 4–6% off the trend line. Fig. 10 shows the benefit of combining both darker and brighter sites for radiometric calibration. It also clearly shows the differences in gain settings between the two sensors. The ETM+ gain settings used here (all high mode) are similar to TM gain settings in bands 1 and 5 but higher than the TM gain settings in bands 2, 3, 4, and 7. This is in accordance with the prelaunch specifications for the ETM+.

5. Cross-calibration results

The M_i slopes derived from the two image pairs (RVPN and NIOB) separately and combined were used in Eq. (11) to generate TM responsivity coefficients (Table 5, Fig. 10). Fig. 11 compares results from the two image pairs with RVPN arbitrarily chosen as the reference case. The consistency between results from the two image pairs varies from negligible differences in spectral band 4 to almost 4% difference in band 7. The average difference is 1.6%, which,
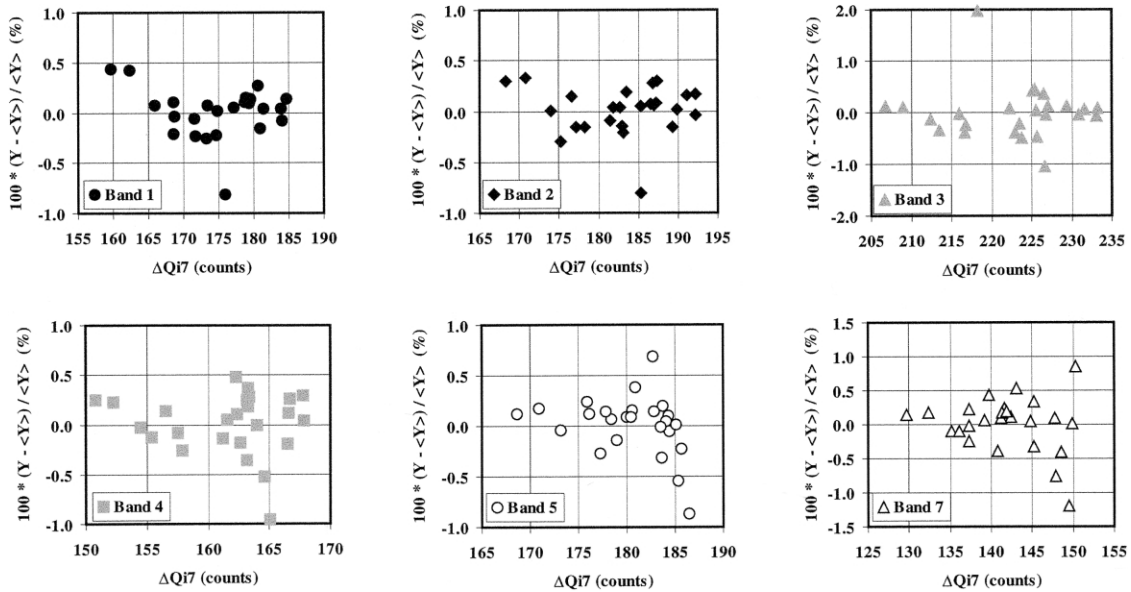
Fig. 9. Residuals between ΔQ_{i5A} values and the zero-intercept linear fit equation for the RVPN subscene pair as a function of ΔQ_{i7} values.

although based on only 12 spectral band cases, is a measure of the repeatability of the cross-calibration approach.

Fig. 11 also shows differences in TM responsivities if spectral band difference adjustments are excluded ($B_i = 1$). Overall, adjustments for spectral band difference appear to be on the order of 2% or less in the visible and near-infrared bands, but greater than that in the shortwave infrared bands. As noted in an earlier part of the paper, the implication is that cross-calibration comparisons that do not benefit from

the surface reflectance spectra and atmospheric optical parameters needed to compute the spectral band difference effect will potentially have an inherent additional uncertainty of several percent.

TM responsivity coefficient updates from 1984 to 1994 are available in the literature based on reflectance-based vicarious calibrations by the UAZ using the calibration site at White Sands, New Mexico (Thome, Markham, Barker, Slater, & Biggar, 1997). In addition, a TM vicarious

Table 5
Slopes (M_i) and sensor responsivities (G_{i5} and G_{i7}) for the two subscenes pairs, including slopes with and without zero intercepts, and for the two subscene pairs combined

Spectral band	M_i (free-intercept)	Intercept (counts)	M_i (zero-intercept)	$100 (1 - R^2)$ (zero-intercept)	G_{i7} (CPUR)	G_{i5} (CPUR)
RVPN subscene data						
1	1.005	1.617	1.014	0.40	1.225	1.243
2	0.5493	0.2945	0.5509	0.45	1.191	0.6561
3	0.5837	1.055	0.5884	2.5	1.538	0.9050
4	0.7161	1.207	0.7235	1.2	1.496	1.082
5	1.008	7.011	1.047	1.3	7.589	7.944
7	0.6525	1.944	0.6662	1.2	21.80	14.52
NIOB subscene data						
1	0.9973	2.346	1.030	1.7	1.225	1.261
2	0.5454	1.267	0.5659	0.77	1.191	0.6740
3	0.5374	2.807	0.5812	0.49	1.538	0.8939
4	0.6629	6.189	0.7226	36	1.496	1.081
5	1.036	0.4631	1.040	0.22	7.589	7.891
7	0.6422	0.01241	0.6424	0.061	21.80	14.00
Combined tandem pair subscene data						
1			1.017	0.026	1.225	1.246
2			0.5525	0.039	1.191	0.6580
3			0.5879	0.017	1.538	0.9042
4			0.7233	0.60	1.496	1.082
5			1.045	0.052	7.589	7.931
7			0.6608	0.26	21.80	14.41

The G_{i5} responsivities are based on the zero-intercept slopes and Eq. (12). R = correlation coefficient and CPUR = counts per unit radiance.
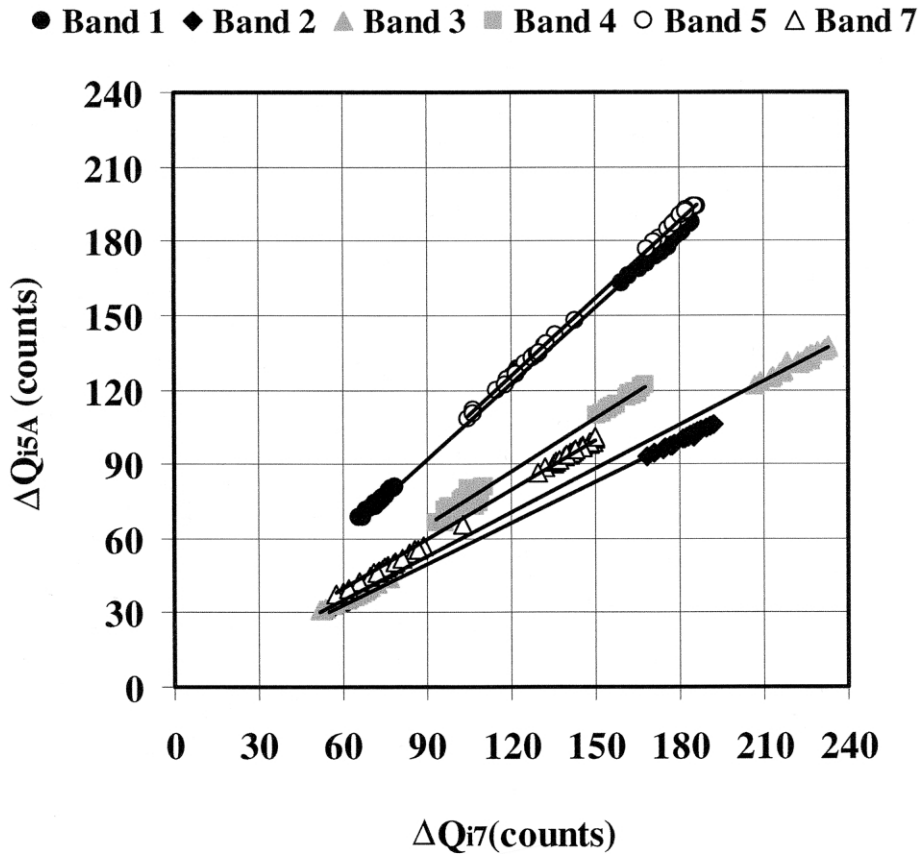
Fig. 10. Plot of grid cell ΔQ_{i5A} and ΔQ_{i7} means combined for the two subscenes. The lines are linear fits with zero-intercepts.
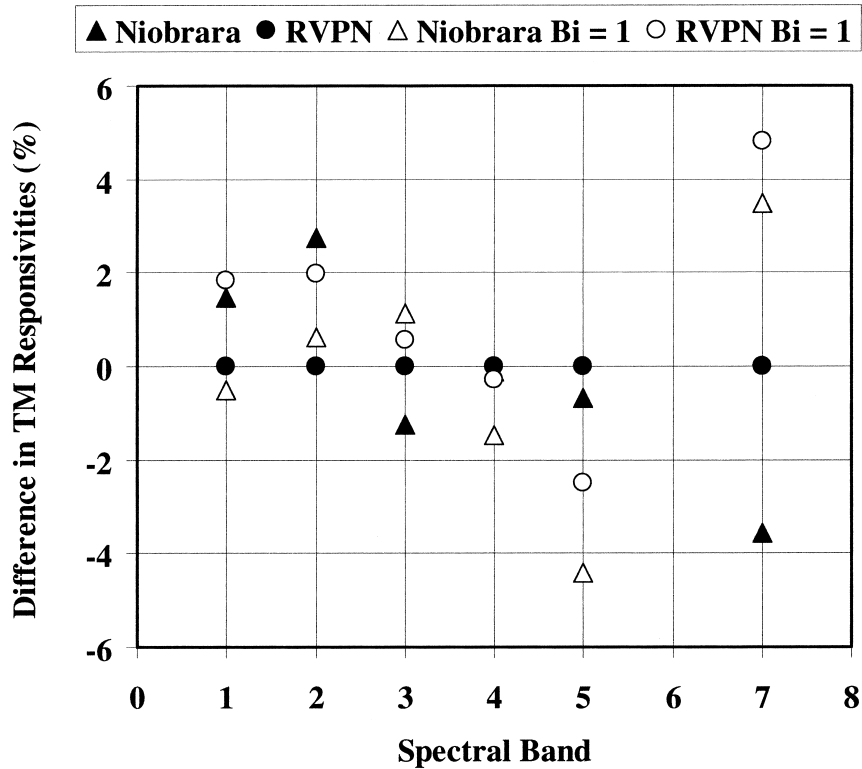

Fig. 11. Comparison of cross-calibration results (G_{i5} responsivities) obtained for the different subscene pairs with respect to the results for RVPN. The open symbols denote results for which no spectral band difference adjustments were made ($B_i = 1$).
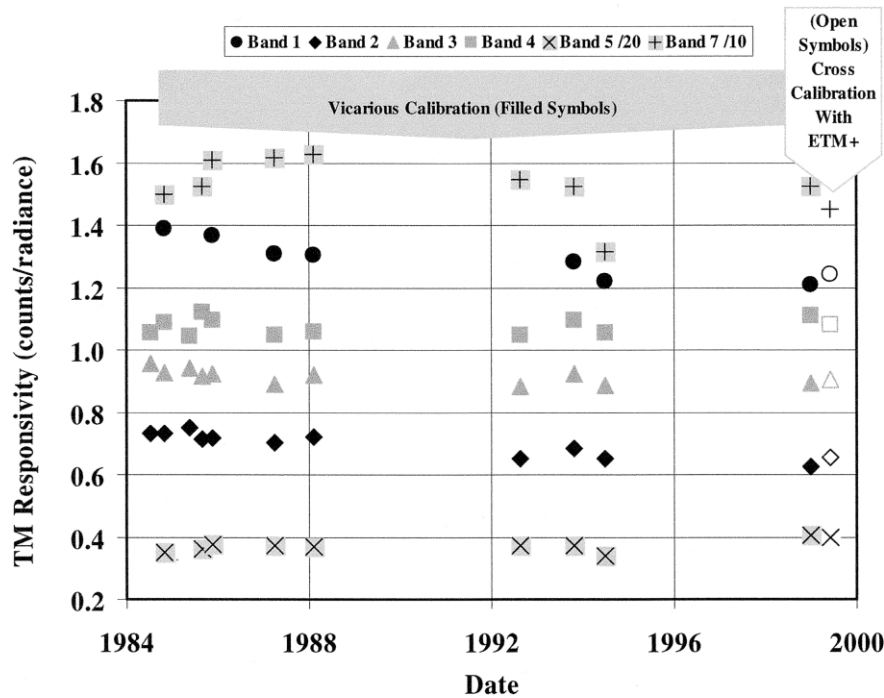
Fig. 12. Tandem-based cross-calibration results (G_{i5} responsivities) from 1999 (larger open symbols) and UAZ vicarious calibrations from 1984–1994. For display purposes TM responsivities in bands 5 and 7 are divided by 20 and 10, respectively.

calibration result for the RVPN test site on June 1, 1999, has kindly been provided by UAZ in advance of publication. These results are plotted in Fig. 12 together with the RVPN results from the present paper based on cross-calibration with respect to the Landsat-7 ETM+ in June 1999, which is a completely independent approach. The average difference between the cross-calibration and vicarious calibration results is 1.8% for bands 1–4 and 3.7% overall, the minimum and maximum differences being 1.1% and 9.3% in spectral bands 3 and 7, respectively (Table 6). Error bars are not shown in Fig. 12, but uncertainties in the vicarious calibration results are reported to be ±5% in the visible and near-infrared bands and approximately 50% greater in the shortwave infrared bands (Thome et al., 1997). The results in the figure are consistent with the possibility that there were no major changes in the Landsat-5 TM responsivity in any of the spectral bands between 1994 and 1999. For spectral bands 5 and 7, it should be noted that there is an additional 5% peak-to-peak variation in the general responsivity trend, likely due to the periodic build up of ice on the window in front of the cold focal plane (Markham, Seiferth, Smid, & Barker, 1998).

SDSU has kindly provided in advance of publication vicarious calibration results for Landsat-5 TM obtained at the NIOB test site on June 2, 1999, when the tandem data were acquired (Black, Helder, & Schiller, 2001). Table 6 presents a comparison of the TM responsivity results from the cross-calibration work and the SDSU vicarious calibration. Agreement to within 3.2% is found in spectral bands 1–3 and 5. SDSU also carried out a vicarious

calibration for the Landsat-7 ETM+ at NIOB on the same day and the ETM+ band 4 result is approximately 10% lower than the nominal ETM+ responsivity provided by the Landsat Project Science Office. This is a possible explanation for the difference between the tandem cross-calibration and SDSU results for TM band 4 (−9.3%). Similarly, the SDSU result for ETM+ band 7 responsivity is approximately 6% higher than the nominal ETM+ value, which could explain part of the 21% difference found between the TM band 7 results. On-orbit calibration updates for spectral bands 5 and 7 have always been characterized by greater uncertainties and so the results in Table 6 for these bands, although unsatisfactory, are not unexpected.

A starting point for estimating the uncertainty of the tandem-based cross-calibration method is the ±3% uncertainty of the ETM+ radiometric calibration (Barker et al., 2000). Additional sources of uncertainty include residual geometric misregistration, small changes in atmospheric conditions between tandem image pair acquisitions, artifacts in the TM image radiometry, and residual uncertainty from spectral band difference adjustments. The jitter analysis indicated a misregistration effect on the order of 0.24% and, although no corroborative analyses have been carried out, experience suggests that the other uncertainties are also well within 1%. If these additional sources of uncertainty amount to a 1–2% effect, the overall root-sum-squared uncertainty for the cross-calibration method is approximately ±3.5%. The near-simultaneity of image acquisition and the similarity of imaging geometry afforded by the tandem configuration are definite advantages in this context.

Table 6
Tandem-based cross-calibration results for Landsat-5 TM (G_{i5} responsivity coefficients) compared to vicarious calibration results obtained by the UAZ and SDSU for the RVPN and NIOB test sites, respectively

Spectral band	1999 RVPN ETM+ cross calibration G_{i5} (CPUR)	1999 UAZ vicarious calibration G_{i5} (CPUR)	Difference relative to cross-calibration G_{i5} (CPUR), %	1999 NIOB ETM+ cross calibration G_{i5} (CPUR)	1999 SDSU vicarious calibration G_{i5} (CPUR)	Difference relative to cross-calibration G_{i5} (CPUR), %	1999 RVPN ETM+ cross calibration G_{i5} (CPUR)	Prelaunch calibration G_{i5} (CPUR)	Difference relative to prelaunch G_{i5} (CPUR), %
1	1.243	1.211	− 2.6	1.261	1.221	− 3.2	1.243	1.555	− 20
2	0.6561	0.6270	− 4.4	0.6740	0.6620	− 1.8	0.6561	0.786	− 17
3	0.9050	0.8953	− 1.1	0.8939	0.9040	1.1	0.9050	1.020	− 11
4	1.082	1.111	2.7	1.081	0.980	− 9.3	1.082	1.082	0.00
5	7.944	8.097	1.9	7.891	7.681	− 2.7	7.944	7.875	0.88
7	14.52	13.17	− 9.3	14.00	16.91	21	14.52	14.77	− 1.7

Percentage differences in these cases are relative to the cross-calibration result. Tandem-based cross-calibration coefficients from RVPN are also compared to prelaunch responsivities, where the percentage difference is with respect to the prelaunch values.

If the spectral signature of the common test site surface is unknown and the spectral band difference effect is 5%, say, then the overall uncertainty approaches ±6%.

Table 6 also compares the 1999 TM calibration update to prelaunch responsivity coefficients for the six solar reflective spectral bands. Differences ranging from − 11% to − 20% exist for spectral bands 1, 2, and 3, whereas the differences are only a few percent or less in spectral bands 4, 5, and 7. Clearly, the use of prelaunch calibration coefficients for the visible bands would lead to significant errors in TOA radiances and any quantities derived from TOA radiances, including surface reflectances retrieved from the TM imagery via atmospheric correction, for example.

6. Concluding remarks

A cross-calibration methodology has been formulated and implemented to use image pairs from the tandem configuration period to radiometrically calibrate the Landsat-5 TM with respect to the Landsat-7 ETM+. The use of large areas common to both the ETM+ and TM image data successfully avoids radiometric effects due to residual image misregistration. The most limiting factor in the approach is the need to adjust for spectral band differences between the two sensors, which requires knowledge about the spectral content of the scene. The spectral band difference effect was found to be more dependent on the surface reflectance spectrum than on atmospheric and illumination conditions. In particular, targets such as dense vegetation and clear water should not be used for radiometric cross-calibration unless surface reflectance spectra are available. Rangeland, grassland, sand, playa, and snow are potentially good candidates for radiometric cross-calibration in the visible and near-infrared bands even in the absence of surface reflectance spectra. However, cross-calibration in the two shortwave infrared bands remains a problem without surface reflectance spectra.

The tandem-based cross-calibration takes advantage of the excellent radiometric performance of the Landsat-7 ETM+ and provides an update to the Landsat-5 TM calibration in the solar reflective bands. Initial trials of the approach with two different tandem image pairs yielded repeatable results for TM responsivity coefficients. For spectral bands 1–3 and likely band 4, the tandem cross-calibration results compare closely to those obtained using independent methods. Additional work is needed to reduce the disagreement in results for the two shortwave infrared bands.

The long-term consistency of the Landsat data record relies heavily on the best efforts and cooperation of several agencies and universities for success. The user community deserves to have a consistent Landsat data record as soon as possible and the success of Landsat-7 is an opportunity to achieve this goal. The present study indicates that the tandem cross-calibration approach can provide a valuable "contemporary" calibration update for Landsat-5 TM based on the excellent radiometric performance of Landsat-7 ETM+. Once retrospective studies have been incorporated to establish a TM calibration record over its mission lifetime to date (Markham et al., 1998), an effort will have to be made to specify and implement algorithms for the proper calibration of archived raw TM data and, wherever possible, existing processed TM data sets.

References

Barker, J. L., Dolan, S. K., Sabelhaus, P. A., Williams, D. L., Irons, J. R., Markham, B. L., Bolek, J. T., Scott, S. S., Thompson, R. J., Rapp, J. J., Arvidson, T. J., Kane, J. F., & Storey, J. C. (2000). Landsat-7 mission and early results. In: *Proceedings of SPIE Conference 3870, Europto SPIE Conference on Sensors, Systems, and Next-Generation Satellites V, Florence, Italy* (pp. 299–311).

Black, S. E., Helder, D. L., & Schiller, S. J. (2001). Irradiance-based cross calibration of Landsat-5 and Landsat-7 Thematic Mapper sensors. *International Journal of Remote Sensing* (in review).

Dinguirard, M., & Slater, P. N. (1999). Calibration of space-multispectral imaging sensors: a review. *Remote Sensing of Environment*, *68* (3), 194–205.

Markham, B. L., Seiferth, J. C., Smid, J., & Barker, B. L. (1998). Lifetime responsivity behavior of the Landsat-5 Thematic Mapper. In: *Proceedings of SPIE Conference 3427, San Diego, CA* (pp. 420–431).

Metzler, M. D., & Malila, W. A. (1985). Characterization and comparison of Landsat-4 and Landsat-5 Thematic Mapper data. *Photogrammetric Engineering and Remote Sensing*, *51* (9), 1315–1330.

O'Neill, N. T., Royer, A., & Nguyen, N. (1996). Scientific and technical report on the development of a modified version of the H5S code which incorporates major features of the 6S code. *CARTEL Technical Report, 1996-020*, 62 pages, Centre d'Applications et de Recherches en Télédétection (CARTEL), Université de Sherbrooke, Sherbrooke, Québec.

Price, J. C. (1989). Calibration comparison for the Landsat 4 and 5 Multi-spectral Scanners and Thematic Mappers. *Applied Optics*, *28* (3), 465–471.

Teillet, P. M., Fedosejevs, G., Gauthier, R. P., O'Neill, N. T., Thome, K. J., Biggar, S. F., Ripley, H., & Meygret, A. (2001). A generalized approach to the vicarious calibration of multiple Earth observation sensors using hyperspectral data. *Remote Sensing of Environment*, *77* (3), 304–327.

Thome, K., Markham, B., Barker, J., Slater, P., & Biggar, S. (1997). Radiometric calibration of Landsat. *Photogrammetric Engineering and Remote Sensing*, *63* (7), 853–858.