Remote Sensing of Environment 191 (2017) 246–257

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Derivation and validation of the stray light correction algorithm for the thermal infrared sensor onboard Landsat 8

Aaron Gerace *, Matthew Montanaro

Center for Imaging Science, Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, NY 14623, United States

ARTICLE INFO

Article history:
Received 6 September 2016
Received in revised form 10 January 2017
Accepted 25 January 2017
Available online 3 February 2017

Keywords:
Landsat 8
TIRS
Stray light
Inter-calibration
Terra/MODIS
Split-window

ABSTRACT

It has been known and documented that the Thermal Infrared Sensor (TIRS) on-board Landsat 8 suffers from a significant stray light problem (Reuter et al., 2015; Montanaro et al., 2014a). The issue appears both as a non-uniform banding artifact across Earth scenes and as a varying absolute radiometric calibration error. A correction algorithm proposed by Montanaro et al. (2015) demonstrated great potential towards removing most of the stray light effects from TIRS image data. It has since been refined and will be implemented operationally into the Landsat Product Generation System in early 2017. The algorithm is trained using near-coincident thermal data (i.e., Terra/MODIS) to develop per-detector functional relationships between incident out-of-field radiance and additional (stray light) signal on the TIRS detectors. Once trained, the functional relationships are used to estimate and remove the stray light signal on a per-detector basis from a scene of interest.

The details of the operational stray light correction algorithm are presented here along with validation studies that demonstrate the effectiveness of the algorithm in removing the stray light artifacts over a stressing range of Landsat/TIRS scene conditions. Results show that the magnitude of the banding artifact is reduced by half on average over the current (uncorrected) product and that the absolute radiometric error is reduced to approximately 0.5% in both spectral bands on average (well below the 2% requirement). All studies presented here indicate that the implementation of the stray light algorithm will lead to greatly improved performance of the TIRS instrument, for both spectral bands.

1. Introduction

The Thermal Infrared Sensor (TIRS) onboard Landsat 8 was designed to continue broadband, long wave infrared measurements of the Earth for the Landsat program (Reuter et al. (2015)). The sensor gathers imagery of the Earth from three detector arrays arranged in a pushbroom configuration. Spectral filters placed on the detector arrays are designed to generate image data in two spectral channels centered on 10.9 and 12.0 μm (known as Band 10 and Band 11, respectively). The detectors along with a four element f/1.64 telescope produce a 15-degree field-of-view (185 km swath width) as Landsat 8 travels in its 705 km altitude orbit.

After launch in early 2013, on-orbit commissioning procedures were performed to characterize the performance of the instrument and adjust the calibration to ensure the radiometric integrity of the data products. Although TIRS exhibited excellent stability and noise characteristics (Montanaro et al., 2014b), two residual unwanted effects in the form of non-uniform banding and a varying absolute radiometric error could not be corrected (Montanaro et al., 2014a; Montanaro et al., 2014c; Barsi et al. 2014; Montanaro et al., 2014d). The magnitude of these effects appeared to be correlated to the magnitude of the radiance outside the field-of-view of the instrument. The ensuing investigation determined that TIRS was plagued by a stray light problem in which out-of-field radiance was entering the optical system and adding a non-uniform signal to the focal plane detectors. With a detailed understanding of the problem, a correction algorithm that leverages an optical stray light model and real-time scene data was developed to adaptively correct the stray light artifacts on a scene-by-scene basis (Montanaro et al., 2015). The algorithm has been refined and validated and has now been approved by the USGS for implementation into the Landsat ground system to produce corrected scenes for users. The implication of its implementation is that TIRS will finally be fully operational for both Bands 10 & 11. The details of the algorithm and validation are presented in this document.

2. Background

After on-orbit checkout of the instrument, it became apparent that each detector was contaminated by an unwanted signal in addition to the direct line-of-sight signal expected from each detector. This additional signal is spatially variant across the field-of-view (FOV) and can

* Corresponding author.
E-mail address: gerace@cis.rit.edu (A. Gerace).

http://dx.doi.org/10.1016/j.rse.2017.01.029

be in excess of 8% of the expected signal from the instantaneous field-of-view (IFOV) of the detector (Montanaro et al., 2014a).

In Earth scene imagery from TIRS, the spatially variant effect manifests itself in two ways. The first effect is a non-uniform banding artifact that appears as a low frequency variation in signal across the FOV. The banding is especially noticeable in areas of a scene that are expected to be uniform (see Fig. 1). Notice from the contrast enhanced image on the right that the banding is seen mostly as a sudden shift in signal between the three focal plane arrays. In terms of science requirements, the resulting banding artifact can be higher than 2% in terms of radiance (vs. the 0.5% requirement). This banding behavior varies by scene and varies over time so that standard flat-fielding methods using the on-board calibrator are not successful when applied to Earth imagery (Montanaro et al., 2014a; Montanaro et al., 2014c).

In addition to the banding problem, the second residual effect indicative of a stray light issue is a time-varying absolute calibration error. As with previous Landsat thermal instruments, TIRS relies on vicarious calibration techniques to characterize the absolute radiometric accuracy of the image products (Barsi et al., 2014). The detected signal of a well-characterized target on Earth (e.g., buoys) is compared to the known temperature of the target propagated to the top of atmosphere (TOA) through a radiative transfer model. Ideally the difference between the detected signal and the modeled TOA radiance would be zero. However, TIRS experiences an absolute radiometric error that is time-varying by season and by location on the Earth. In general, the error is greater when the area surrounding the scene is warm and lower when the surrounding area is cool. The resulting absolute radiometric error can be as high as 4–5% in Band 10 and 8–9% in Band 11 (vs. the 2% requirement). Additionally, the large variance in the radiometric accuracy of TIRS (sigma = 0.12 $W/m^2/sr/\mu m$ for Band 10 and 0.20 $W/m^2/sr/\mu m$ for Band 11) remains even when the average error is removed (Barsi et al., 2014). As with the banding artifact, the varying absolute radiometric error cannot be corrected for with conventional means, i.e., a gain and bias correction. The current Landsat 8 L1 T product available to users for the thermal bands (Band 10 and 11) has the average absolute error removed (0.29 $W/m^2/sr/\mu m$ for Band 10 and 0.51 $W/m^2/sr/\mu m$ for Band 11), (Barsi et al., 2014; Montanaro et al., 2014d).

The large uncertainty in the absolute radiometric accuracy as well as the non-uniform banding could not be removed. The fact that the magnitude of both of these issues were correlated with the magnitude of the out-of-field radiance were explained by the stray light hypothesis in which out-of-field radiance (stray light) adds a spatially-varying signal to the focal plane. That non-uniform stray light signal appears as the banding artifact while the varying magnitude of the stray light signal appears as the large variance in the radiometric uncertainty.

The stray light geometry was definitively confirmed with a series of targeted lunar scans in which the moon was positioned in multiple locations outside of the FOV of the instrument. At certain locations in the out-of-field (primarily at approximately 13 degrees from the optical axis), signal well above the background noise was detectable which provided a clear demonstration of the stray light problem (Montanaro et al., 2014a). The lunar scan data along with laboratory measurements of the spare TIRS telescope assembly provided enough information on scattering angles to facilitate the construction of an optical model of the instrument. A reverse ray trace through the optical model was performed for every detector element on the focal plane, which closely matched the lunar and laboratory measurements. Referring to Fig. 2, the ray traces are essentially stray light maps that describe the locations (angles) and the magnitudes (weights) of the stray light source for every detector element. Each detector stray light map is unique to that detector since the stray light effect varies with location on the focal plane (Montanaro et al., 2015). These per-detector stray light maps are used as the basis of the stray light correction algorithm.

3. Algorithm description

An algorithm, which attempts to remove the stray light artifacts from TIRS Earth imagery, must be able to significantly reduce the banding artifacts as well as reduce the overall absolute calibration uncertainty. Additionally, it must be flexible enough to perform well on every Earth scene regardless of the scene content and be able to execute fairly quickly in order to prevent a severe processing backlog. The details and initial results of such an algorithm were demonstrated by Montanaro et al. (2015). The algorithm has been refined since it was originally proposed to the Landsat Calibration team by incorporating lessons learned from the initial iteration and training of the algorithm. A step-by-step description of the current algorithm is presented here.

3.1. Theoretical basis

The stray light removal algorithm assumes that there is a functional relationship between the out-of-field radiance and the magnitude of the

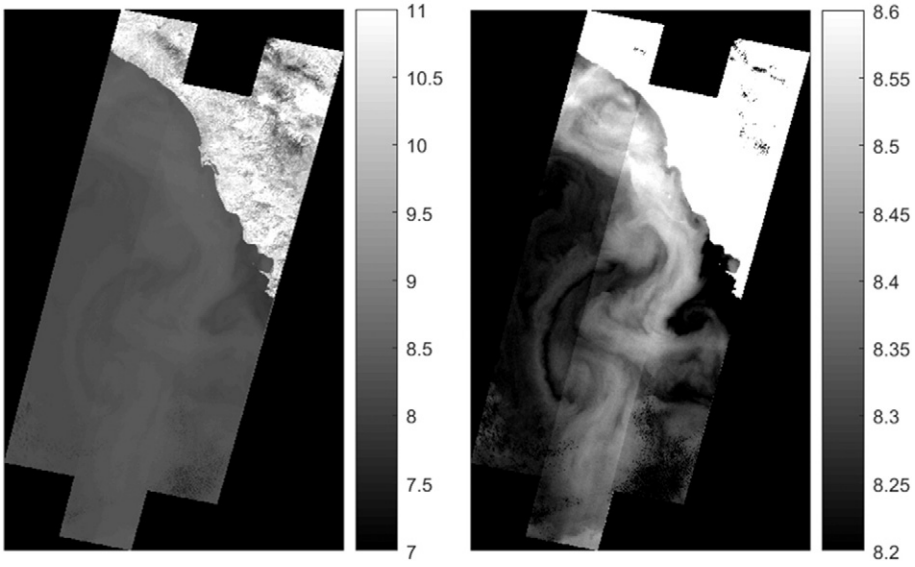


Fig. 1. (Left) TIRS Band 10 image of the coast of Baja California (WRS path 040) and (right) a contrast-enhanced image (units of spectral radiance). The banding artifacts are clearly visible as sudden shifts in pixel values between the three focal plane arrays.

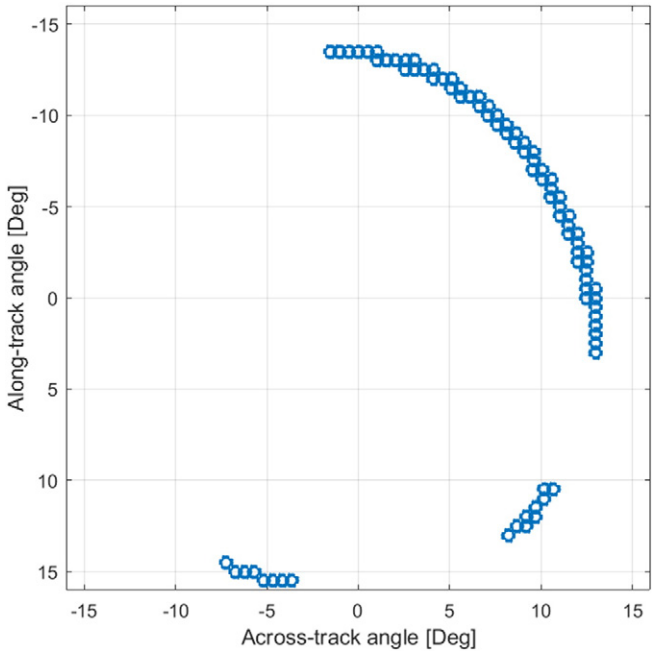


Fig. 2. A stray light map developed from the TIRS optical model for an arbitrary detector on the TIRS focal plane. Each circle indicates a location in the out-of-field relative to the optical axis of the instrument that contributes to additional signal on the focal plane. Note that each of the 1920 detectors per band of TIRS has its own unique stray light map.

extra signal on the detectors (see Section 3.3). The recorded signal by a detector element j is equal to the direct signal from its IFOV plus the signal scattered from the out-of-field (i.e., the stray light signal):

$$total\_signal_j = direct\_signal_j + scattered\_signal_j \quad (1)$$

It is proposed here that the 'true' *direct signal* for each detector element may be recovered from its *total signal* if the *scattered signal* can be estimated. The stray light maps produced by the reverse ray traces through the TIRS optical model provide line-of-sight locations of the stray light scattering relative to the optical axis. If the radiance level at these out-of-field locations can be estimated, then the scattered signal onto a detector j is proportional to the summation of these out-of-field radiances such that,

$$scattered\_signal_j = f\left[\sum_i w_{j,i} \cdot L_{j,i}\right] \quad (2)$$

where $L_{j,i}$ is the out-of-field radiance for detector j originating from location i in the stray light map, $w_{j,i}$ is the relative weighting factor associated with location i, and f indicates a functional relationship. Once the scattered signal component is calculated for every detector and every line of a TIRS image, it can be subtracted from the TIRS image to remove the stray light signal. In other words, a stray light 'image' is calculated that represents the stray light signal for every pixel in the TIRS image and is subtracted from the standard L1 T product to produce a corrected product.

3.2. Estimation of out-of-field radiance

The key challenge for the stray light correction algorithm as described in Section 3.1 is determining the out-of-field radiances for every scene as it is collected by TIRS. In an ideal scenario, out-of-field thermal data would be available from a wide-field-of-view sensor that obtains data coincident with TIRS and in similar spectral channels and a similar spatial resolution. The detector stray light maps would be utilized to sample the locations in the wide-field sensor data and calculate the magnitude of the scattered light signal on each detector (Eq. (2)), (Montanaro et al., 2015).

Unfortunately, such a dataset does not exist for all TIRS Earth scenes. It was previously proposed and demonstrated that a geostationary sensor such as the GOES imager (Band 4) could be utilized to obtain the out-of-field radiances for TIRS scenes over the mid-latitude Americas (Montanaro et al., 2015). However, the large differences in view angle, spectral band shape, and spatial resolutions between GOES and TIRS were found to introduce significant uncertainty into the sampled out-of-field radiance values of GOES. In addition, GOES data is not available for high-latitudes and for locations outside of the Americas. Other wide-field instruments would have to be utilized for those locations, which introduces additional concerns of cross-calibration and archiving of such image data.

To avoid the limitations associated with utilizing image data from an external wide-field sensor to estimate out-of-field radiance, Montanaro et al. (2015) proposed the novel concept of utilizing the TIRS image data itself as a surrogate for the out-of-field knowledge. The optical stray light maps are used along with location and pointing telemetry from the spacecraft to calculate where on the Earth (latitude/longitude) the scattering sources are located for a given detector at a given time. As illustrated in Fig. 3, some of the scattering originates from locations along the TIRS swath while the remaining scattering originates from outside the ± 7.5 degree field-of-view of the instrument. For scattering locations that fall within the TIRS swath, the radiance values for these locations are sampled directly from the TIRS data. For scattering locations that fall outside the swath where no TIRS data is available, the algorithm samples the pixel values at the edge of the TIRS swath and assumes that these edge values are a reasonable estimate of the radiance values outside the field-of-view (Montanaro et al., 2015). In this way, the algorithm utilizes TIRS scene data itself as an estimate of the out-of-field data. The algorithm is sometimes referred to as the "TIRS-on-TIRS" algorithm by the Landsat Calibration team due to its self-reflexive nature.

Once the out-of-field radiance has been estimated for each scattering location, it is weighted according to the optical map data (i.e., some line-of-sight locations result in more scattering than other locations) and then summed. The summed value, from the right side of Eq. (2), represents the total weighted out-of-field scattered light value as described by the reverse ray traces from the optical model for a given detector at a particular location in an Earth scene. The summed value for a detector is *not* in terms of radiance since the weighting functions in Eq. (2) are not scaled in an absolute sense to the total scattered signal on the detector. Therefore, the summed values must be adjusted back into terms of radiance by comparing them to the actual scattered signal from a reference dataset (see next section).

3.3. Derivation of a functional relationship

A regression process is performed as a one-time initialization step to derive the per-detector functional relationship between the summed out-of-field value and the magnitude of the stray light signal on the detector in order to scale the summed value to units of radiance. To determine this relationship (i.e., the functional relationship in Eq. (2)), the detector's measured signal is compared to a known "truth" value to determine the stray light on that detector (Montanaro et al., 2015). A sufficient source of truth data over the entire FOV of the instrument is not readily available since image data from the current Landsat 8 orbit is unique in that it does not spatially and temporally overlap another known image dataset globally. However, shortly after launch the Landsat 8 spacecraft under-flew the Terra spacecraft for several days such that data from the TIRS and MODIS sensors were nearly coincident (within roughly 20 min of each other) and the fields-of-view of the two sensors were centered on each other. Since Bands 31 & 32 of MODIS are spectrally similar to Bands 10 & 11 on TIRS, respectively, and

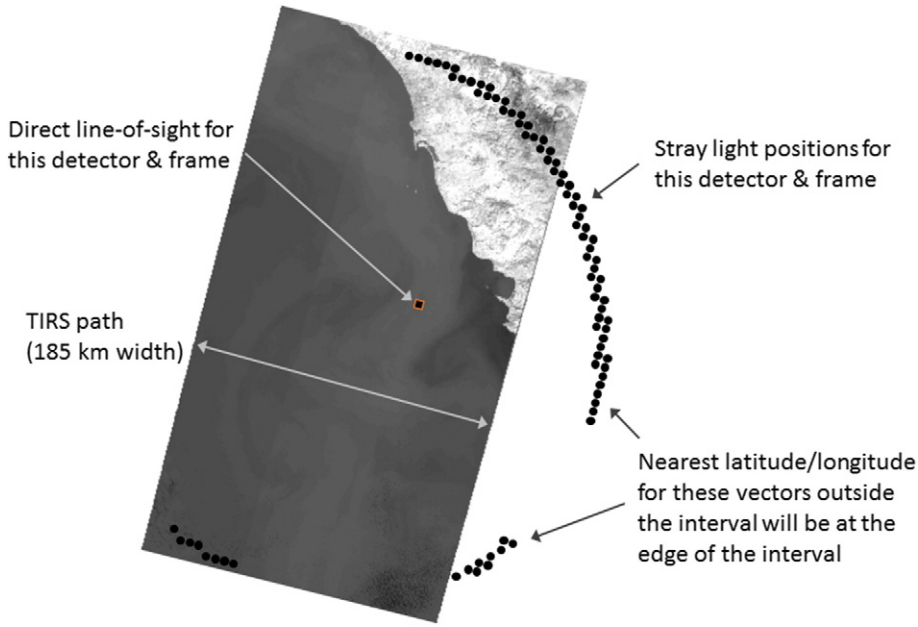


Fig. 3. Illustration of the reverse ray traces from the optical stray light map for a single detector. Some of the stray light source locations fall within the TIRS scene interval while others fall outside the field-of-view. For the locations that fall within the FOV, the TIRS scene radiance is sampled. For stray light locations that fall outside the FOV where there is no scene data, the pixel values at the edge of the TIRS scene are used as a surrogate for the radiance distribution outside the field-of-view. This assumes that the edge pixel values roughly correlate to the radiance field outside the FOV.

considering the Terra/MODIS measured performance is significantly better than its original design goal of 0.5% absolute radiance, (Xiong et al., 2015), its thermal data was treated as truth for TIRS. Once TIRS data is resampled and spatially registered to MODIS data, the difference between the two datasets yields an estimate of the magnitude of the stray light signal on the TIRS detectors.

The algorithm is "trained" on a per-detector basis by regressing the stray light signal (from comparisons to MODIS) and the integrated out-of-field value (determined from the reverse ray traces and TIRS sampling methodology described in Section 3.2). Eq. (2) then becomes,

$$\mathrm{TIRS}_j - \mathrm{MODIS}_j = f\left[\sum_i w_{j,i} \cdot L_{j,i}\right], \tag{3}$$

with the *scattered signal* from Eq. (2) calculated as the difference between TIRS and MODIS during the Landsat 8/Terra under-flight period. For the training step, fully resolved water pixels provided the best regression data since offshore water bodies are spatially uniform and invariant in the thermal infrared over the 20 min between TIRS and MODIS acquisitions. Additionally, snow and ice scenes were included to provide regression data at low radiance values. The out-of-field radiance for each detector in these scenes was calculated via the "TIRS-on-TIRS" methodology while the stray light signal was calculated from the difference between TIRS and MODIS. By determining the stray light signal and integrated out-of-field values over a range of Earth scene temperatures (approximately 200 K to 300 K), it was found that the relationship (in Eq. (3)) between the two variables was approximately linear as illustrated for an arbitrary detector in Fig. 4. The coefficients of the best-fit line are calculated for each detector and, once found, they are fixed and the algorithm considered trained. The stray light maps (angles and weights) along with the linear regression coefficients are unique for each detector. In the correction process described in the next section, each detector is treated independently.

3.4. Correction procedure

Once the relationship between the stray light signal and the summed out-of-field value is adequately developed, the algorithm is ready to be applied operationally to data. For each of the 700+ TIRS scenes collected in a given day, the stray light correction algorithm will be applied by sampling the TIRS scene data according to the stray light maps and the TIRS-on-TIRS sampling methodology, weighting & summing the samples, applying the linear regression coefficients derived in Section 3.3, and subtracting the calculated

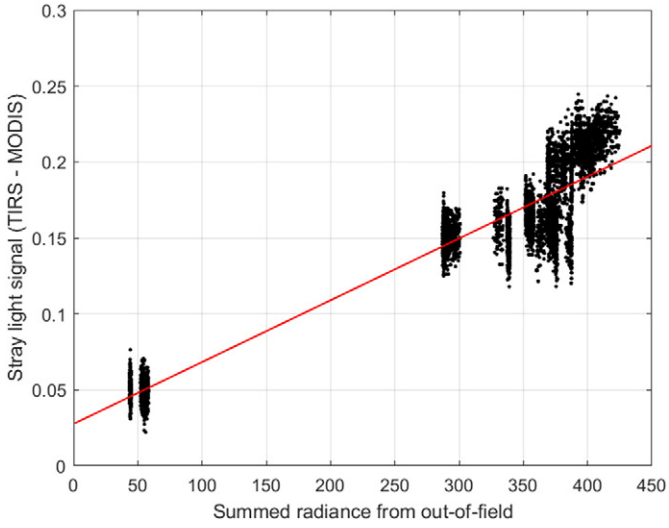


Fig. 4. Scatter-plot of stray light signal radiance vs. summed out-of-field value using the TIRS-on-TIRS methodology for an arbitrary detector. Data points represent values from seven reference scenes. The best fit linear regression is represented by the red dotted line.

stray light signal from the TIRS image data. The algorithm takes on the final form,

$$corrected\_signal_j = total\_signal_j - \left\{ \alpha_j \left[\sum_i w_{j,i} \cdot L_{j,i} \right] + \beta_j \right\}, \quad (4)$$

where α_j and β_j are the linear regression coefficients for each detector j, $L_{j,i}$ is the radiance from the ith scattering location from the optical model (sampled from TIRS data) for detector j, and $w_{j,i}$ is the weighting factor from the ith location for detector j. To summarize, for any given Earth scene, the algorithm procedure is as follows:

1) Calculate the out-of-field radiance for a given detector by sampling the TIRS scene according to the optical stray light map for the detector via the TIRS-on-TIRS methodology described in Section 3.2, apply the weighting factors to the sampled radiances, and sum all the weighted values.
2) Apply the linear coefficients to the summed out-of-field values to estimate the magnitude of the stray light radiance signal on the detector.
3) Repeat (1) and (2) for every detector and every line in the TIRS scene to produce an image of the stray light magnitude for every pixel.
4) Subtract the stray light image from the uncorrected TIRS scene image to create the final corrected image product.

The U.S. Geological Survey (USGS) has implemented the stray light correction algorithm into an off line version of the Landsat Product Generation System allowing a convenient way to quickly produce corrected TIRS Earth scenes. Recalling the example from Figs. 1 and 5 (left) shows the scene before the stray light correction (i.e., the product available to users) while Fig. 5 (right) shows the scene after applying the stray light correction algorithm. Note that both images are contrast-enhanced to the same scale. Visually, in this example, the non-uniform banding artifact has been significantly reduced upon applying the stray light correction algorithm.

4. Validation

In order to quantitatively characterize its performance, a series of independent scenes not used in the regression training were run through the stray light correction algorithm and compared to the standard product that is currently available to users. For a quantitative assessment of the differences between the standard product and the stray light corrected product, the scenes were compared to calibrated Terra/MODIS data, which was treated as truth in this study. Recall that using Terra/MODIS as truth for this study was considered valid as the MODIS Bands 31/32 match TIRS spectrally and the Terra/MODIS measured performance is significantly better than its original design goal of 0.5% absolute radiance, (Xiong et al., 2015). Note that although the ideal correction should result in image data that is the same as truth, this study seeks to assess the performance of the stray light correction algorithm with respect to the current TIRS product available to users. Ideally, to be considered for implementation into Landsat's Product Generation System, the stray light corrected product will be consistently closer to truth than the current product for both bands of TIRS.

4.1. Scene selection

The selection of scenes to perform a proper assessment of the proposed algorithm was challenging due to the scarcity of calibrated space-borne thermal data available as truth. The ideal truth source would be an instrument with similar sensor characteristics (i.e., spatial and spectral resolution), collected at the same time, and through the same atmosphere as TIRS. While a geosynchronous imager (e.g., GOES or MeteoSat) was initially considered, data from such instruments are of limited value due to their coarse spatial resolution and significantly different view angles through the atmosphere. Point measurements of ground targets, such as those from water buoys, provide a convenient source of truth for well-behaved instruments but do not enable an assessment of the banding reduction over the entire image. Although the Terra/MODIS instrument was initially dismissed since its orbit is out of phase with

Landsat 8 (i.e., TIRS images are at the extreme edge of the FOV of Terra/MODIS), it was reconsidered for the Landsat 8 and Terra underflight period previously discussed.

Recall from Section 3.3 that the TIRS FOV was centered on the MODIS FOV during the under-flight period so that scenes collected during this time period were attractive due to the minimal impact of view angle effects as both MODIS and TIRS imaged through similar atmospheric

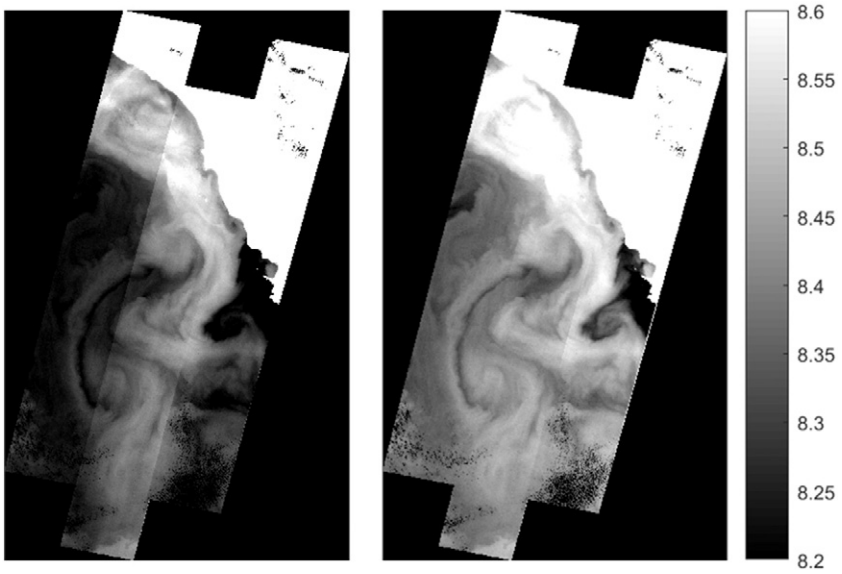


Fig. 5. TIRS Band 10 image of the coast of Baja California (WRS path 040) before (left) and after (right) stray light correction. Visually there is a reduction of the banding artifact over the current (uncorrected) image product.

paths. Although MODIS Bands 31 & 32 are spectrally similar to TIRS Bands 10 & 11, a small adjustment must still be made to the MODIS data to compensate for the small sensor passband shape differences with TIRS. There are, however, two potential limitations of using MODIS data from the under-flight period. The first is the time offset between the Terra and Landsat 8 spacecraft. Any Earth phenomena that change within the roughly 20-min offset between the two sensors will introduce an artificial difference when the MODIS and TIRS data are compared. Clouds are the main concern since they may move a significant distance in 20 min. The second potential issue is the difference in spatial resolution between the two sensors. TIRS data (at 100 m) must be resampled to MODIS resolution (at 1 km) in order to register the two datasets for comparison. Highly non-uniform Earth scenes will result in large differences between TIRS and MODIS due to the spatial resampling. Typical land scenes are usually too non-uniform (i.e., contain high spatial frequencies) where the spatial resampling introduces artificial variability. Therefore, for this analysis, mostly uniform water scenes with adequate cloud free regions were chosen to minimize errors in the truth data and to isolate residual errors in the stray light correction process. Typically, offshore water scenes are spatially uniform at thermal infrared wavelengths and the high thermal inertia of water minimizes temperature fluctuations over the 20 min offset.

Of the approximately 1000 TIRS under-flight scenes collected over the 3-day period, only about 20 were suitable for this assessment. The ideal scene to challenge the TIRS-on-TIRS sampling assumption should have clouds in the out-of-field and/or complicated land/water boundaries. Uniform scenes (i.e., scenes with land both in and out of the FOV or scenes with water in and out of the FOV) were avoided for this analysis as were scenes with significant cloud contamination. Note that the scenes used in the training of the algorithm were also excluded from this validation process. The locations marked in Fig. 6 indicate the World Reference System 2 (WRS-2) path/rows that were chosen from the under-flight period to assess the performance of the stray light removal algorithm. Each location is shaded based on the average apparent temperature observed over the water for that scene. Note by the color bar that the chosen scenes span the middle portion of the TIRS required operating temperature range, [240 K–330 K]. Fig. 6 highlights the potential attractiveness of the algorithm's ability to operate on any scene without external data as it enables a worldwide solution.

4.2. Algorithm performance characterization

As previously mentioned, to quantify the potential efficacy of the stray light correction algorithm against the current Landsat 8 TIRS product, band-shape adjusted Terra/MODIS Band 31 & 32 data were used as truth for TIRS Band 10 & 11 data, respectively. The standard TIRS products for all scenes highlighted in Fig. 6 were acquired from USGS and the same scenes were run through the stray light correction algorithm. Difference images between MODIS and TIRS were then generated for both the stray light removed data and the standard product. Note that although the TIRS and MODIS products are in units of spectral radiance, the validation comparisons are presented here in apparent temperature (units of Kelvin) in order to frame the results in units that typical thermal infrared users might be more comfortable with.

Fig. 7 shows an example of the resulting difference images for Path 010, Row 030 off the east coast of the United States for both Bands 10 (top) & 11 (bottom). The images at the upper left shows the difference between the current product and MODIS and the corresponding difference between the stray light corrected product and MODIS. The standard product exhibits strong non-uniform banding and has a significant absolute error (approximately 1–2 K for Band 10 and 4 K for Band 11). The stray light corrected product, on the other hand, exhibits dramatically reduced banding as well as substantially lower absolute error when compared to MODIS. To quantify the magnitude of the difference of the two TIRS products relative to truth, two profiles through the scene were obtained as illustrated at the right hand side in Fig. 7. As alluded to previously, care was taken to avoid obtaining profiles over the clouds that are evident in the northeast portion of the scene. Since the TIRS data are being compared to MODIS, most of the spatial variation in the Earth scene should be taken into account when differenced with MODIS so that

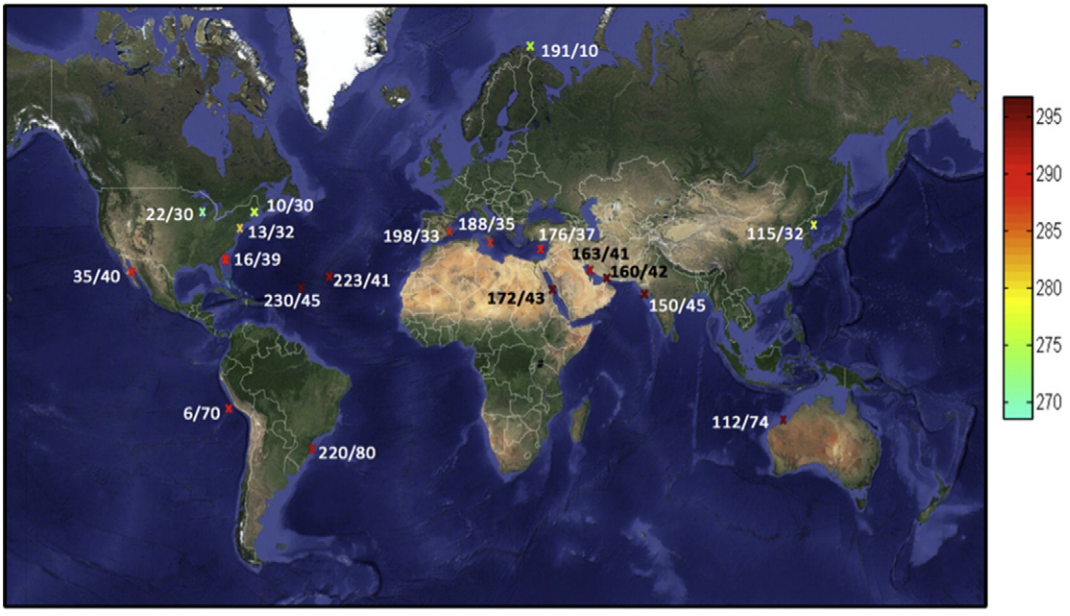
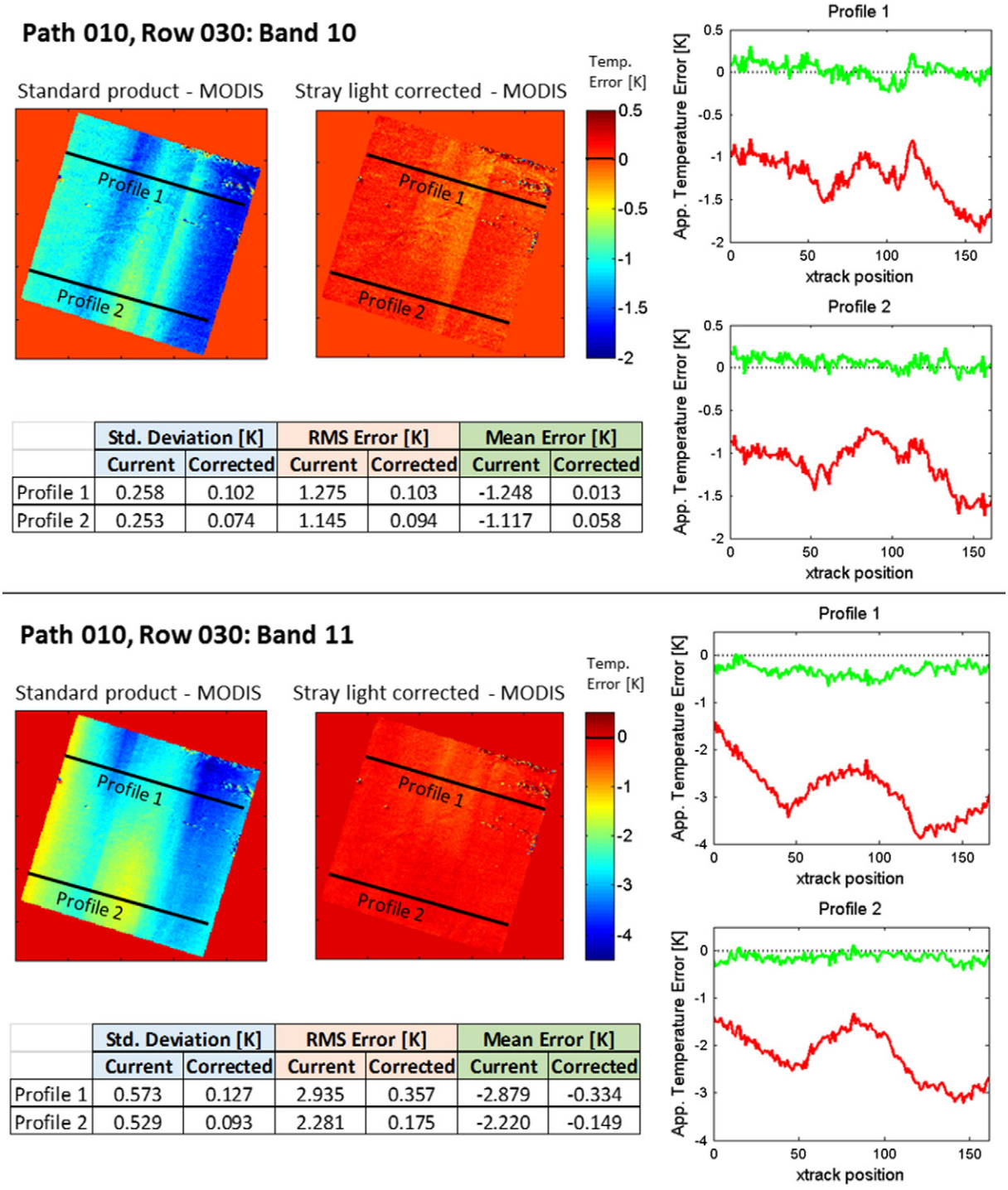


Fig. 6. The marked locations indicate the WRS-2 path/rows that were chosen from the under-flight period to assess the performance of the stray light removal algorithm. Each location is shaded based on the average apparent temperature observed over the water for that scene. The color bar is in units of Kelvin.



	Std. Deviation [K]		RMS Error [K]		Mean Error [K]	
	Current	Corrected	Current	Corrected	Current	Corrected
Profile 1	0.258	0.102	1.275	0.103	-1.248	0.013
Profile 2	0.253	0.074	1.145	0.094	-1.117	0.058

	Std. Deviation [K]		RMS Error [K]		Mean Error [K]	
	Current	Corrected	Current	Corrected	Current	Corrected
Profile 1	0.573	0.127	2.935	0.357	-2.879	-0.334
Profile 2	0.529	0.093	2.281	0.175	-2.220	-0.149

Fig. 7. Comparison of a TIRS Earth scene off the coast of the United States (Path 010, Row 030) before and after the stray light correction for Band 10 (top) and Band 11 (bottom). The images are differences between the TIRS standard product and MODIS (left) and the stray light corrected product and MODIS (right). Profiles in the scenes are graphed on the right for the standard product (red) and the stray light corrected product (green). The tables list the statistics calculated on the profiles. The data displayed here illustrate the effectiveness of the stray light correction algorithm.

ideally the difference profile would be flat and nearly zero. The profiles indicate that the stray light correction (green profiles) significantly reduces the non-uniformity and the absolute error in the TIRS scene over the current product (red profiles). Additionally, standard deviation, root-mean-square (RMS) error, and mean absolute error statistics are calculated on the scene profiles and are listed in the tables in Fig. 7. As indicated in the tables, the standard deviation, RMS, and mean error are significantly reduced as a result of the stray light correction algorithm.

The majority of the scenes highlighted in Fig. 6 show a similar reduction of error as described above for Path 010, Row 030 shown in detail in Fig. 7. The scenes have a variety of out-of-field geometries in order to test the ability of the algorithm to estimate out-of-field radiance (recall that the radiances at the edge of a TIRS scene are

utilized as a surrogate for the out-of-field radiance). Fig. 8 shows the difference images between the TIRS current product and MODIS and the stray light corrected product and MODIS for a subset of the validation scenes. Note that multiple WRS-2 rows of data were stitched together to generate these mosaics and that the land was masked in order to avoid the large spatial and temporal differences due to the varying spatial resolutions and the time offset between TIRS and MODIS as mentioned previously. A visual inspection of these four cases shows a clear and dramatic reduction in error when the TIRS-on-TIRS correction algorithm is applied relative to the standard calibration method.

4.3. Summary statistics

Thirty-five profiles were collected from the difference images associated with the scenes highlighted in Fig. 6 in the same manner described in Section 4.2. Profiles were also obtained from difference images generated for a desert scene (Path 198, Row 046) and a snow/ice scene (Path 191, Row 010) to draw attention to the complexities associated with performing validation over land. The standard deviations of all thirty-five profiles were obtained to characterize the spatial variability of the data against truth and are summarized in Fig. 9. The standard deviation in units of Kelvin is reduced by roughly half, on average, over the current product indicating that the stray light corrected product offers a more spatially uniform product. There are a couple of profiles where the standard deviation did not reduce by the same magnitude as the other profiles or the standard deviation became slightly higher after the correction. These cases can be easily explained by the truth data as described earlier and are not the result of the stray light correction (e.g., clouds or temperature variations in the water introduce large differences between MODIS and TIRS skewing the results).

As discussed in Section 2, the current product available to users does not meet the banding and radiometric accuracy requirements due to stray light contamination. To quantify the impact of the stray light correction algorithm with respect to the TIRS requirements, the RMS error for all the profiles is summarized in Fig. 10 in terms of a percent radiance error. This was accomplished by dividing the RMS error (in terms of radiance) for each profile by the mean value of the profile (in radiance). Similar to the standard deviations, the RMS error has been drastically reduced after applying the stray light correction so that the average residual RMS error is approximately 0.5% for both bands (recall that the required absolute error limit for TIRS is 2%). The implication of this reduced error is apparent as it indicates that TIRS may potentially become a fully functioning, dual-band instrument with an absolute accuracy that is in-line with the Terra/MODIS requirement by implementing this stray light correction algorithm. Note that this algorithm does not include any typical radiometric calibration (e.g., flat-fielding based on the internal calibration source or scene statistics) so there is potential to further enhance its radiometric accuracy in the future.

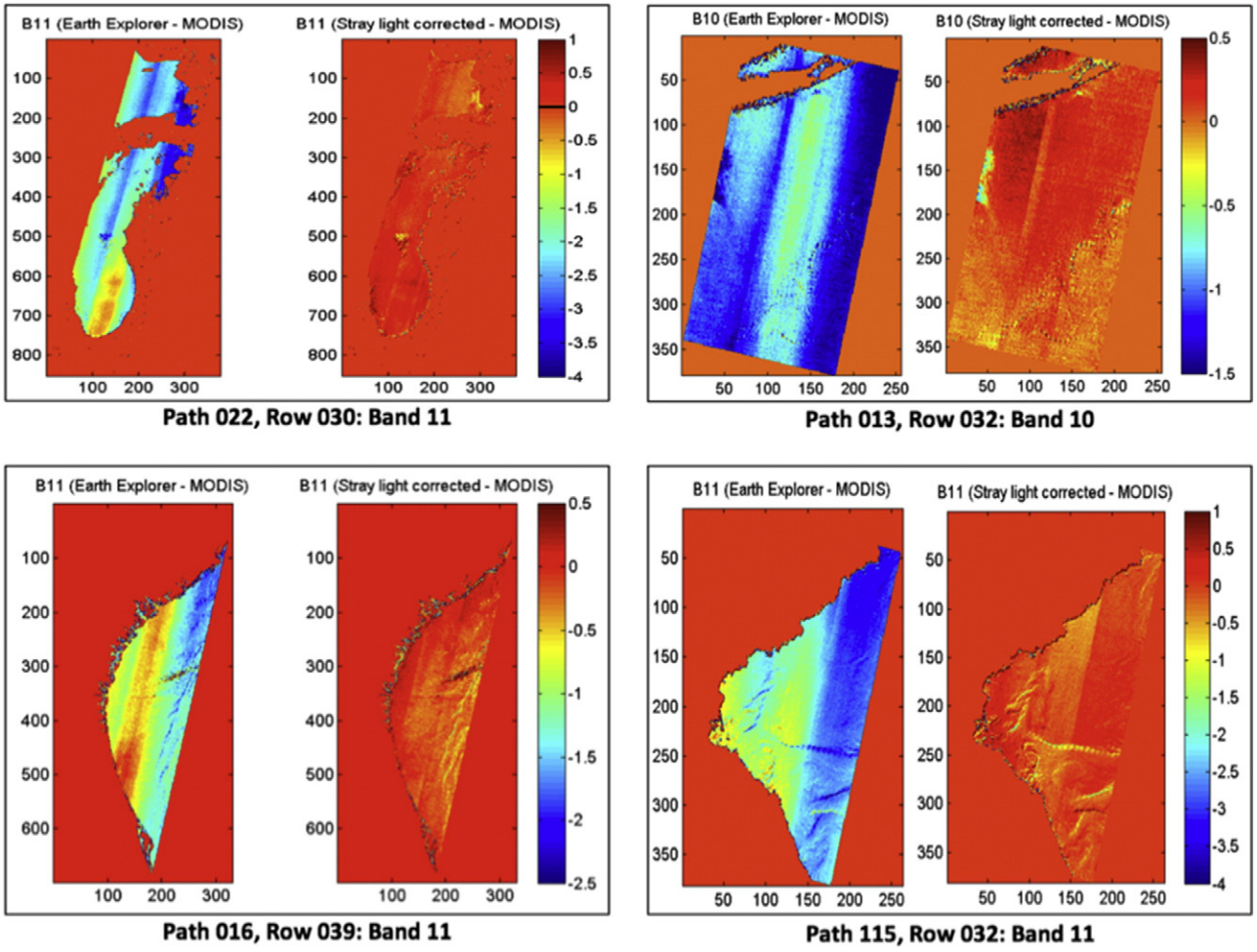


Fig. 8. Apparent temperature difference images in units of Kelvin for a subset of the validation scenes identified in Fig. 6. Note that multiple WRS-2 rows of data were stitched together to generate these mosaics and that the land was masked.

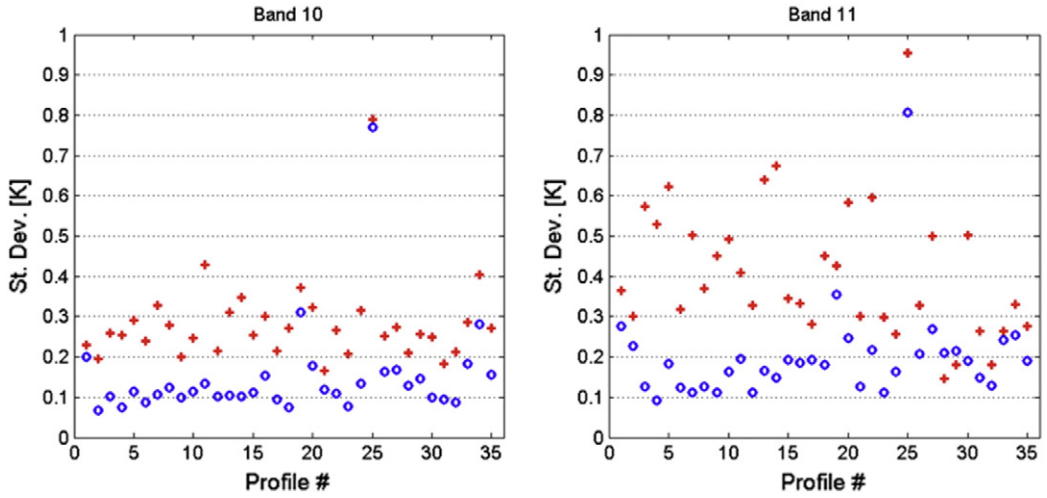


Fig. 9. Summary of the standard deviation of the MODIS-subtracted scene profiles of the current TIRS product (red pluses) and the stray light corrected product (blue circles) for TIRS Band 10 (left) and Band 11 (right). Units are in apparent temperature [K].

It should be noted in Fig. 10 (Band 11) that the error associated with four profiles appear out-of-family from the other data values (indicated by the enclosed data points in the figure). These four profiles are the land scenes (shown in Fig. 11) and bring attention to the complexities associated with using another space-borne thermal dataset over land (i.e., Terra/MODIS) as truth. The two profiles from the top scene (snow/ice) in Fig. 11 correspond to the errors enclosed in the square in Fig. 10. The behavior of these data points is likely caused by the spatial resampling of TIRS data (at 100 m) to MODIS data (at 1 km), which introduces differences between the TIRS and MODIS products that are not associated with stray light and skew the analysis. Similarly, the two profiles from the bottom scene (desert) in Fig. 11 correspond to the errors enclosed in the circle in Fig. 10. The desert sand in the scene heats significantly under solar insolation between the MODIS and TIRS acquisition times thereby introducing a temperature bias in the RMS calculation, which causes the RMS error to increase for this particular case. Again, these out-of-family values are believed to be a result of the truth data and not an artifact of the stray light algorithm.

4.4. Implications of the stray light correction algorithm on split-window atmospheric correction

Considering the instrument's nominal dual-band design, a desire to use the split-window atmospheric correction method for retrieving surface temperature from TIRS data is present in the remote sensing community. This is evident based on several published journal articles where authors apply the split-window technique to currently available TIRS data despite recommendations to the contrary from USGS EarthExplorer, (Kamran et al., 2015; Yu et al., 2014; Nikam et al., 2016; Tan et al., 2016). While ongoing work seeks to assess the potential for the split-window technique to be used with TIRS data, a preliminary

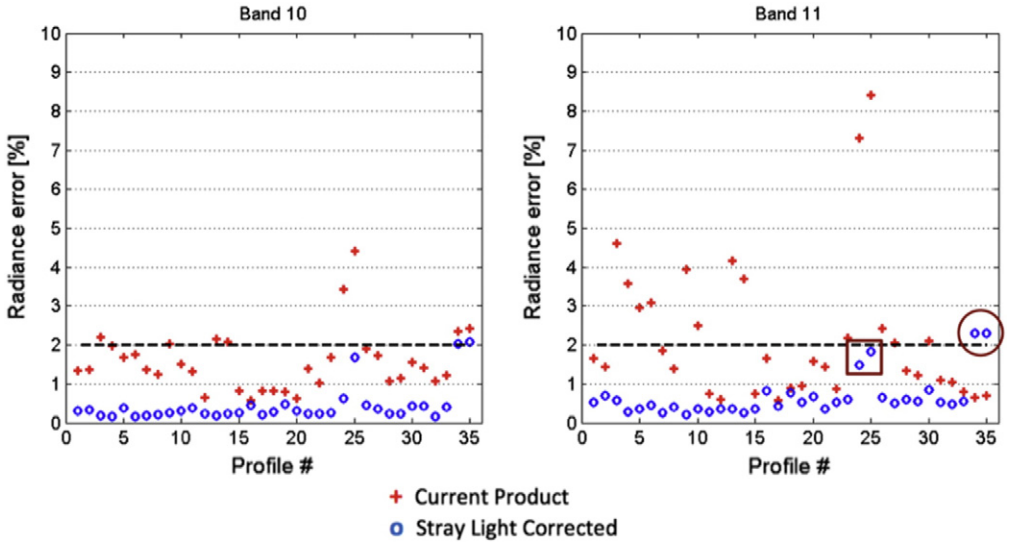


Fig. 10. Summary of the RMS error of the MODIS-subtracted scene profiles for the current TIRS product (red pluses) and the stray light corrected product (blue circles) for TIRS Band 10 (left) and Band 11 (right). Units are in percent of the average scene radiance and the dotted line indicates the requirement limit. Note that Profiles 34 and 35 are of desert sand in which the sand heated up significantly in the time between the MODIS and TIRS acquisitions. Profiles 24 and 25 are of snow/ice and are susceptible to spatial resampling artifacts in the MODIS truth data.

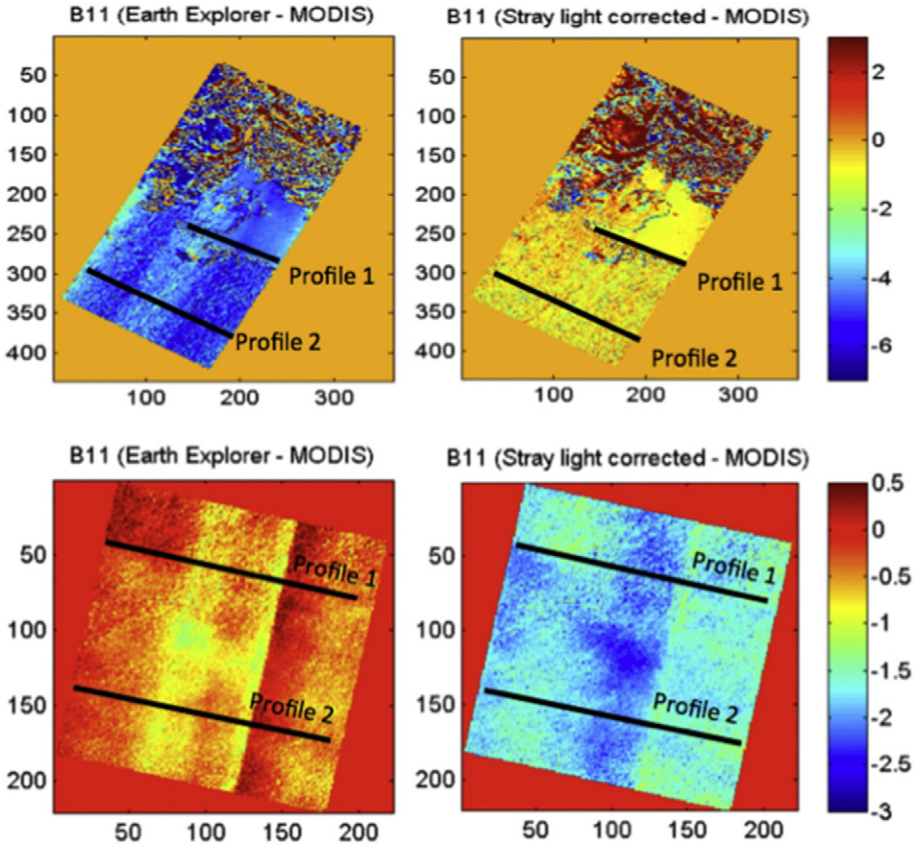


Fig. 11. Profiles obtained over difference images for a snow/ice scene (Path 191, Row 010) (top) and a desert scene (Path 198, Row 046) (bottom), drawing attention to the complexities associated with performing validation over land. Difference images are in apparent temperature [K].

study is presented here to motivate the impact of the stray light correction algorithm on surface temperature retrieval accuracy.

A full treatment of the split-window technique is beyond the scope of this article. However, Eq. (5) shows that the land surface temperature (LST) can be derived from the apparent temperature in two bands, denoted T_i and T_j, assuming the average emissivity (ε) and the difference in emissivity ($\Delta\varepsilon$) between the two bands are known, as well as the column water vapor, (Du et al. (2015)). Note that Du et al. (2015) conduct a model-based study that derives b_i coefficients based on column water vapor for Landsat 8 TIRS bands.

$$\mathrm{LST} = b_0 + \left(b_1 + b_2\frac{1-\varepsilon}{\varepsilon} + b_3\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_i+T_j}{2} + \left(b_4 + b_5\frac{1-\varepsilon}{\varepsilon} + b_6\frac{\Delta\varepsilon}{\varepsilon^2}\right)\frac{T_i+T_j}{2} + b_7(T_i-T_j)^2 \quad (5)$$

To assess the potential utility of TIRS with respect to the split-window technique, both the TIRS standard product and the TIRS stray light corrected product are obtained for Path 13, Row 32 from Fig. 6. Recall that this data was collected during the under-flight so the corresponding Terra/MODIS data is again utilized as truth in this study, i.e., due to its high radiometric fidelity for Bands 31 & 32, Terra/MODIS represents a best-case scenario with respect to LST retrieval. For simplicity, only the emissivity of water is used here, as this scene is comprised primary of water. As no ancillary knowledge of column water vapor (CWV) was obtained for this study, the b_i coefficients associated with the entire possible CWV range is used in Eq. (5), as suggested by Du et al. (2015). While CWV measurements would be desirable in practice to obtain as accurate an absolute retrieval of LST as possible, this study seeks to measure the feasibility of using the current product and the stray light corrected product relative to a calibrated sensor such as MODIS. Accordingly, Eq. (5) is applied to all three products and difference images (Fig. 12) were obtained in the same fashion as described in Section 4.2. The left image in Fig. 12 shows the difference in retrieved LST for the current product and MODIS (denoted $\mathrm{LST_{CP}} - \mathrm{LST_M}$) while the right image shows the difference in retrieved LST for the stray light corrected product and MODIS (denoted $\mathrm{LST_{SLC}} - \mathrm{LST_M}$).

Visually, the stray light corrected product generates a LST product that is significantly closer to the MODIS LST product than the current TIRS data available to users. Temperature errors of over 6 K in the water have been reduced to 0–1 K for most of the scene. To quantify the impact of the stray light correction algorithm on LST retrieval across the full field-of-view of TIRS, profiles were obtained over the scene as shown in Fig. 12 (left) and plotted in Fig. 12 (right). Notice for Profile 1 that temperature errors of 1–5 K observed in the current product (red plot) are reduced to less than 1 K upon applying the stray light correction algorithm (green plot). Profile 2 was obtained across the Gulf Stream boundary, representing a challenging scenario with regards to validation for reasons discussed earlier, i.e., moving water over the 20 min temporal offset between data collects will contaminate the truth data. Even in this scenario, errors of 2–5 K observed in the current product (red plot) are reduced to 0–2 K upon applying the stray light correction algorithm (green plot). Notice that the dip in the green plot for Profile 2 corresponds to the boundary in the difference image where the cold ocean water transitions to the warm Gulf Stream water. If we ignore this location in the profile and consider the truth data contaminated there, the remaining portion of Profile 2 contains

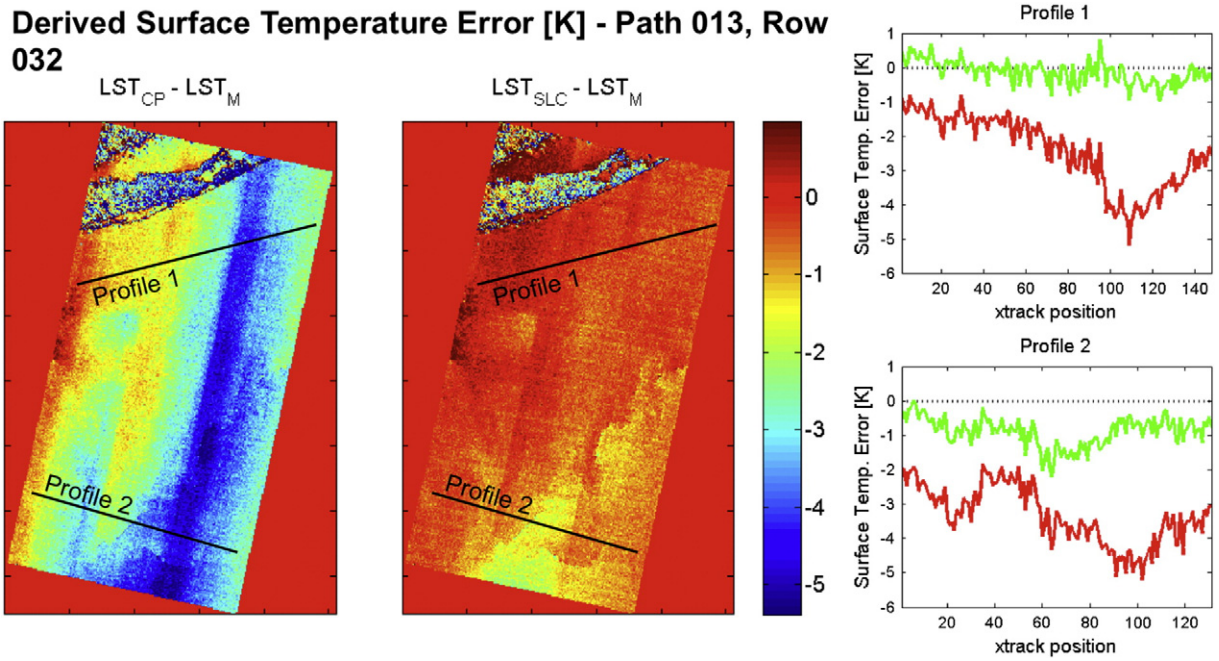


Fig. 12. Error in derived surface temperature from the split window technique between TIRS (current product) and MODIS (left figure) and TIRS (stray light corrected product) and MODIS (right figure). The profiles indicated in the figures are plotted on the right for the current product (red) and the stray light corrected product (green).

errors of less than 1 K for the stray light corrected product. Albeit preliminary, this investigation illustrates the renewed potential for TIRS to be used for split-window atmospheric correction.

5. Conclusions

Although the noise and stability performance of the Thermal Infrared Sensor has been excellent while operating on-orbit, persistent image artifacts in the form of banding and absolute radiometric error pointed to a significant stray light problem with the instrument. A software solution to the problem was designed that takes advantage of an optical stray light model of the instrument to estimate the magnitude of the stray light signal for every pixel in a TIRS Earth scene. The optical maps for each detector are projected onto the Earth's surface and expressed as a series of latitude/longitude points to sample the TIRS image data. The sampled values from this "TIRS-on-TIRS" methodology are weighted according to their relative importance and summed and scaled into an estimate of the additional stray light radiance for the detector. The stray light signal is then subtracted from the Earth image to remove the stray light artifacts in the image.

The stray light correction algorithm significantly reduces the banding and absolute error artifacts from TIRS imagery of the Earth to make it more consistent with legacy Landsat thermal band instruments. Not only are the banding artifacts visually reduced in Earth scene imagery, but analyses of the scene profiles before and after correction quantitatively show that the banding is reduced from 2% or higher in terms of radiance before correction to under the 0.5% requirement after the correction. Similarly, the absolute radiometric accuracy is reduced from as high as 9% in radiance (Band 11) to approximately 0.5% in both bands (versus the 2% requirement). Additionally, the algorithm does not require any additional ancillary data as input, which allows it to function on any Earth scene and run in near real time. These characteristics have enabled the algorithm to be approved for implementation into the USGS Landsat Product Generation System starting in early 2017.

All validation efforts to date indicate that the implementation of the stray light correction algorithm presented here will enhance an underperforming Band 10 of TIRS and revive its non-operational Band 11. The implications of applying this algorithm are exciting in that it will not only enhance radiative-transfer based LST retrievals that are commonly applied to broadband instruments but will also satisfy a desire by the remote sensing community to take advantage of a split-window based approach. As with all instruments in the Landsat program, the performance of the TIRS sensor even after implementation of the stray light correction will continue to be monitored to ensure the expected radiometric accuracy for all users.

Acknowledgements

The authors would like to thank NASA Goddard (NNX16AB56G), specifically Jim Irons and Brian Markham, for providing the funding to make this effort possible. Additionally, the authors would like to thank Tim Beckmann of USGS for implementing the stray light correction algorithm into the off-line version of the Landsat Product Generation System. His effort significantly facilitated the validation studies presented here.

References

Barsi, J., Schott, J., Hook, S., Raqueno, N., Markham, B., Radocinski, R., 2014. Landsat-8 thermal infrared sensor TIRS vicarious radiometric calibration. Remote Sens. 6 (11), 11607–11626.

Du, C., Ren, H., Qin, Q., Meng, J., Zhao, S., 2015. A practical split-window algorithm for estimating land surface temperature from Landsat 8 data. Remote Sens. 7, 647–665.

Kamran, K., Pirnazar, M., Bansouleh, V., 2015. Land surface temperature retrieval from Landsat 8 TIRS: comparison between split window algorithm and SEBAL method. Proc. SPIE 9535, Third International Conference on Remote Sensing and Geoinformation of the Environment. 9535-03.

Montanaro, M., Barsi, J., Lunsford, A., Rohrbach, S., Markham, B., 2014a. Performance of the Thermal Infrared Sensor on-board Landsat 8 over the first year on-orbit. Proceedings of SPIE, Earth Observing Systems XIX. 9218-17.

Montanaro, M., Lunsford, A., Tesfaye, Z., Wenny, B., Reuter, D., 2014b. Radiometric calibration methodology of the Landsat 8 thermal infrared sensor. Remote Sens. 6 (9), 8803–8821.

Montanaro, M., Gerace, A., Lunsford, A., Reuter, D., 2014c. Stray light artifacts in imagery from the Landsat 8 thermal infrared sensor. Remote Sens. 6 (11), 10435–10456.

Montanaro, M., Levy, R., Markham, B., 2014d. On-orbit radiometric performance of the Landsat 8 thermal infrared sensor. Remote Sens. 6 (12), 11753–11769.

Montanaro, M., Gerace, A., Rohrbach, S., 2015. Toward an operational stray light correction for the Landsat 8 thermal infrared sensor. Appl. Opt. 54 (13), 3963–3978.

Nikam, B., Ibragimov, F., Chouksey, A., Vaibhav, G., 2016. Retrieval of land surface temperature from Landsat 8 TIRS for the command area of Mula irrigation project. Environ. Earth Sci. 75 (1169), 1–17.

Reuter, D., Richardson, C., Pellerano, F., Irons, J., Allen, R., Anderson, M., Jhabvala, M., Lunsford, A., Montanaro, M., Smith, R., Tesfaye, Z., Thome, K., 2015. The Thermal Infrared Sensor (TIRS) on Landsat 8: design overview and pre-launch characterization. Remote Sens. 7 (1), 1135–1153.

Tan, K., Liao, Z., Du, P., Wu, L., 2016. Land surface temperature retrieval from Landsat 8 and validation with geosensor network. Front. Earth Sci. 1–15.

Xiong, X., Wu, A., Wenny, B., Madhavan, S., Wang, Z., Li, Y., Chen, N., Barnes, W., Salomonson, V., 2015. Terra and Aqua MODIS thermal emissive bands on-orbit calibration and performance. IEEE Trans. Geosci. Remote Sens. 53 (10), 5709–5721.

Yu, X., Guo, X., Wu, Z., 2014. Land surface temperature retrieval from Landsat 8 TIRS-comparison between radiative transfer equation-based method, split window algorithm and single channel method. Remote Sens. 6, 9829–9852.