remote sensing

MDPI

Article

Assessment of Spatial Characterization Metrics for On-Orbit Performance of Landsat 8 and 9 Thermal Infrared Sensors

S. Eftekharzadeh Kay [1,*], B. N. Wenny [1], K. J. Thome [2], M. Yarahmadi [1], D. J. Lampkin [3], M. H. Tahersima [1] and N. Voskanian [2]

[1] Science Systems & Applications, Inc., 10210 Greenbelt Rd, Lanham, MD 20706, USA
[2] Sciences and Exploration Directorate, NASA GSFC, Greenbelt, MD 20771, USA
[3] Booz Allen Hamilton, Laurel, MD 20707, USA
* Correspondence: sareh.eftekharzadeh@nasa.gov

Abstract: The two near-identical pushbroom Thermal Infrared Sensors (TIRS) aboard Landsat 8 and 9 are currently imaging the Earth's surface at 10.9 and 12 microns from similar 705 km altitude, sun-synchronous polar orbits. This work validates the consistency in the imaging data quality, which is vital for harmonization of the data from the two sensors needed for global mapping. The overlapping operation of these two near-identical sensors, launched eight years apart, provides a unique opportunity to assess the sensitivity of the conventionally used metrics to any unexpectedly found nuanced differences in their spatial performance caused by variety of factors. Our study evaluates spatial quality metrics for bands 10 and 11 from 2022, the first complete year during which both TIRS instruments have been operational. The assessment relies on the straight-knife-edge technique, also known as the Edge Method. The study focuses on comparing the consistency and stability of eight separate spatial metrics derived from four separate water–desert boundary scenes. Desert coastal scenes were selected for their high thermal contrast in both the along- and across-track directions with respect to the platforms ground tracks. The analysis makes use of the 30 m upsampled TIRS images. The results show that the Landsat 8 and Landsat 9 TIRS spatial performance are both meeting the spatial performance requirements of the Landsat program, and that the two sensors are consistent and nearly identical in both across- and along-track directions. Better agreement, both with time and in magnitude, is found for the edge slope and line spread function's full-width at half maximum. The trend of averaged modulation transfer function at Nyquist shows that Landsat 8 TIRS MTF differs more between the along- and across-track scans than that for Landsat 9 TIRS. The across-track MTF is consistently lower than that for the along-track, though the differences are within the scatter seen in the results due to the use of the natural edges.

Keywords: spatial calibration; thermal remote sensing; edge method; modulation transfer function; TIRS instrument; Landsat

check for updates

Citation: Eftekharzadeh Kay, S.; Wenny, B.N.; Thome, K.J.; Yarahmadi, M.; Lampkin, D.J.; Tahersima, M.H.; Voskanian, N. Assessment of Spatial Characterization Metrics for On-Orbit Performance of Landsat 8 and 9 Thermal Infrared Sensors. *Remote Sens.* **2024**, *16*, 3588. https://doi.org/10.3390/rs16193588

Academic Editor: Jose Moreno

Received: 9 August 2024
Revised: 30 August 2024
Accepted: 18 September 2024
Published: 26 September 2024

1. Introduction

The Thermal Infrared Sensor (TIRS) aboard Landsat 8 was deployed on 11 February 2013, and its successor, Landsat 9 TIRS (TIRS-2), joined it on a similar orbit on 27 September 2021. These two thermal sensors have significantly advanced our scientific understanding of Earth's thermal dynamics. TIRS instruments are pivotal for monitoring surface temperatures across the globe, offering precise and consistent thermal data to study climate patterns, environmental changes, and natural resource management. By capturing the emitted thermal radiation from the Earth's surface in two spectral bands, centering at 10.9 and 12 micron, TIRS instruments provide essential information for assessing land surface temperatures, identifying heat anomalies, and detecting variations in urban and rural environments.

Earth's landscapes exhibit intricate thermal patterns, and variations in spatial resolution can significantly influence the precision of temperature measurements. A consistent spatial quality ensures confident comparison of thermal data across time and space, enabling longitudinal studies of environmental changes. Consistent spatial resolution is equally essential for detecting subtle variations and trends during the monitoring of urban heat islands or the assessment of land surface temperatures. Moreover, maintaining a consistent spatial quality in thermal imagery (e.g., between TIRS and TIRS-2) is indispensable for compilation of reliable and comparable datasets that underpin the integrity of Earth observation research and its practical applications and ensures confident trend analysis of environmental factors across the calibrated instruments in different platforms.

The two TIRS instruments are dual channel push-broom sensors with a passively-cooled, four-element refractive telescope that directs the incoming beam onto the focal plane, and its rotating flat mirror points the focal plane to either nadir, the space view port, or to the on-board temperature-controlled calibration blackbody [1,2]. In both TIRS instruments, the optical system is kept at an operational temperature of ∼186 K, and a mechanical cryocooler actively keeps the focal plane below 40 K. Both of their focal planes consist of three 512 × 640 arrays of quantum well infrared photodetectors (QWIPs) to cover their 185 km swaths [1,3]. Both TIRS instruments passed the defined radiometric and spatial system requirements during the pre-launch tests in the Thermal Vacuum Chamber (TVAC). The TIRS-2 instrument aboard Landsat 9 is believed to be fairly similar, yet with some improved features compared to Landsat 8 TIRS [1–3]. An important design difference between the two was devised for TIRS-2 after investigations for the root cause of the calibration errors that were discovered to be caused by the stray light shortly after the launch of Landsat 8 [4]. The stray light correction has been investigated to have a positive impact for atmospheric correction and surface temperature measurements, as it removes the gap between the measured temperatures by both instruments. Since their launch and beginning of their routine operations, independent studies have assessed spatial and radiometric performance of both TIRS instruments [5–11]. This work is an independent characterization of TIRS spatial performance over an overlapping time in their operation.

Motivation and Goals for This Study

Although the logical thread that connects the dedicated sections toward the final conclusions is embedded in the text throughout the paper, a clear statement on motivations, expectations and goals for this study can be an effective reference for the reader.

- **Foundation**

This study is based on the fact that TIRS instruments aboard Landsat 8 and 9 are near-identical in design, calibration, operation, orbital parameters, operating modes, relative spectral response, and transmission profiles of their bands. Cross-comparisons of sensors aboard various platforms is a well-known and practiced approach to radiometric calibration validation, but two participating thermal sensors (with publicly available data) have rarely been as identical in all of these aspects as the TIRS instruments on Landsat 8 and 9. Since their performance degradation over time is carefully monitored by the operating teams of the launched platforms, any remaining differences on the conventionally measured metrics can be as informative on the potentials of the metrics themselves, as they are on the sensor's performance.

- **Motivation**

Compiling trade studies for the more sophisticated instruments of the future motivates a revisit of the commonly used metrics for pre- and post-launch calibration validation. Spatial metrics are generally inherited from the original works at wavelength regime and applied to other regimes. Some of those metrics may be found to not be as sensitive to image quality changes in thermal bands. Replacing less effective approaches or metrics with their more effective counterparts can impact the planning, site selection, and measurement tools for the commissioning periods of the launched platforms. Besides spatial performance comparison of the two identical sensors, a revisit of the spatial metrics for thermal imagery to assess their sensitivity to observing condition and scanning direction. Such metric assessments may prove useful in calibration validation planning of the future thermal sensors. The system requirements for both TIRS instruments were public knowledge, which helped us to, on the one hand, conduct an independent verification of those requirements and ascertain the sanity of our measurement algorithms and, on the other hand, make an assessment of the metric sensitivities for both sensors over the course of one year of their operation (Landsat 8 TIRS's tenth year and Landsat 9 TIRS-2's first full year of operation).

- **Expectations and goals**

  Similarly designed and calibrated sensors on similar orbits and similar operating platforms are expected to behave similarly under ideal and matching operating conditions. But, in practice, small on-board calibrator malfunctions or variations in their response can introduce non-uniformities and trigger false pixel-to-pixel variation that dilute the intrinsic contrast of the observed scenes over the resolved scales by the sensor. These effects ultimately degrade the sharpness of the images, an effect that can be exacerbated by resampling into smaller pixel sizes. The goal was to: (1) Scout the coastal scenes around the globe including the previously selected scenes for edge profile measurement, and establish a data set of co-observed scenes with as similarly low cloud coverage as possible. (2) Generate the trends of the same set of metrics on the same observed scenes by both sensors and ideally only days apart. The measured metrics trends in this work have 1-sigma uncertainty envelopes around them to facilitate significance assessment of potential upward or downward trends if they persist for three or more months. (3) Identify the metrics that exhibit more sensitivity to observing conditions or scanning direction (across- or along-track), even in a generally consistent trend for the two metrics. And (4) whether a consistent behavior was confirmed or not, lay out conclusions on the metric sensitivity to a change in observing condition or scanning direction, and/or the metrics ability on assessing the image quality (e.g., sharpness of the edges) and its robustness to bypass various observing conditions and yield values that consistently correlates with image quality. The sensitivity loss of spatial metrics as they are applied to images at longer wavelength is not surprising, and a natural extension of what is presented in metric trends of TIRS instruments was a commentary on which one of the commonly measured spatial metrics retain (or lose) their sensitivity to observing condition and edge sharpness at 10–12 microns. Equally worthy of note are the confluence of analysis and considerations that categorized some of the nuanced metric differences as *noticeable but not significant*. Given the generally higher quality of Landsat imagery in all bands, compared to contemporary thermal sensors with public access, hints at the limits of the on-orbit data and/or Edge Method in resampled thermal images and validating pre-launch assessments.

This paper is structured as follows: Section 2 describes the scene selection criteria and data acquisition and lays out the measurement methods and metric definitions; Sections 3 and 4 are dedicated to the results and discussions, and Section 5 delivers a summary of the work and conclusions. All of the discussed metrics in this work have been measured for all 12 of the qualified images of listed sites, at both B10 and B11 (10.9 and 12 μm) bands. "AT" and "XT" indicate edge orientation in along and across the sensor's scanning track.

2. Materials and Methods

2.1. Data

Spatial performance evaluation of satellite imagery for thermal sensors via the knife-edge technique, also known as the Edge Method or Tilted Edge Method, has been successfully used in validating the spatial quality of satellite images across a wide range of channels on various platforms (e.g., [12–18]). This technique relies on high contrast thermal boundaries (edges) to characterize the transition between the two sides of the edge. The edge targets shall be chosen such that they best represent the resolving power of the imaging system. In addition to the sharp thermal transition, this method relies on the thermal uniformity on either sides of the edge. Consistently high contrast between the two sides and thermal stability of these targets are more desirable for long-term monitoring of the spatial performance of the sensors than thermal events such as volcanic eruptions.

Geometrically defined features, such as bridges and harbors (with enough thermal contrast with surrounding waters), provide lasting structural uniformity along the edge that the naturally formed edges (such as cliffs and coastlines) do not provide. With all of their imperfections, the naturally formed thermal edges are used for spatial characterization metrics for their high thermal contrasts, diversity of locations and orientations with respect to the orbital tracks of various platforms. Orbital motion of the platforms adds directional property to the scanning procedure. It should be noted that, in this work, along-track and across-track terms do not refer to scanning style differences (i.e., scanning from top to bottom or from side to side), but refer to the orientation of the edge image in the scanned scene. A north–south oriented edge image is used to characterize the across-track edge response and a perpendicular edge image to across-track direction, is used for along-track edge response characterization.

Table 1 lists our chosen edge targets in this work. High contrast edges (such as coastline) in desert areas are reliable candidates for the edge technique, especially those that sustain their thermal contrast for most of the year. Among the qualified coastlines with the aforementioned criteria, those with accommodating orientation for the near-polar orbits of Landsat 8 and 9 (at 98° inclination) were selected. The SAHA, OMAN, and YMEN sites were chosen by a study on spatial performance of Landsat 8 TIRS [9]. LIBY was also chosen by [5] for post-launch spatial performance assessment of Landsat 9 TIRS-2. Among the longer list of candidate scenes, these sites proved most thermally uniform, and seasonally reliable sites with relatively consistently high contrast coastlines for edge response measurements. The numbers separated by comma in the fifth column are associated with the total number of qualified images for Landsat 9 TIRS and Landsat 8 TIRS, respectively. If no cloud-free image existed in a 30 day interval, the cloud coverage threshold was gradually increased from zero until at least one usable image was found for a $\sim$30-day interval for the two instruments throughout 2022.

Table 1. A summary table of the participating scenes in this work. Only one of images for SAHA from Landsat 8 TIRS and Landsat 9 TIRS2 has between 10 and 20% cloud coverage. Out of 24 images for the LIBY site, only two images from Landsat 8 TIRS have between 10 and 20% cloud coverage. All of the images from OMAN, YMEN, and DUQM have 0% cloud coverage. The edge orientations with respect to the sensor track in all of the listed sites are no larger than 8 degrees.

Site (Acronym)	Type	Path/Row	Edge SNR	Scenes LS9, LS8	Lat., Lon. [deg.]
West Sahara (SAHA)	XT	206/045	30–90	10, 12	(21.45, −17.0)
North Africa (LIBY)	AT	186/038	30–70	12, 12	(31.25, 16.10)
East Oman (OMAN)	XT	158/046	30–65	5, 5	(19.68, 51.71)
South Yemen (YMEN)	AT	163/050	40–75	5, 5	(13.96, 47.86)
Port of Duqm—Oman (DUQM)	AT, XT	158/046	40–70	5, 5	(19.68, 57.71)

Figure 1 depicts the general location of those scenes. OMAN, YMEN, and SAHA have been previously chosen for spatial quality assessment of Landsat 8 TIRS, as these shorelines have near optimal angle of ±8° from the sensors orbital orientation for across-track spatial characterization [9]. LIBY in Northern Africa (Path/Row:186/38) also passed similar physical and positional criteria for along-track spatial performance characterization (shorelines with $8 \pm 8^\circ$ orientation [9]). Note that, even for the shared sites with past studies, no information about the exact sections of the shorelines that were used in their study was available. Our method of moving a measuring window along the shores and finding a section with consistently maximal thermal contrast and edge profile behavior over time was used to select the final sites among the candidate sites.

Figure 1. General locations of the chosen scenes denoted by colored squares. The sizes of the squares do not represent the scaled projections of the tiles. The underlying map is a snapshot of the interactive tile map in earthexplorer.usgs.gov.

The top and bottom left panels of Figure 2 show examples of the thermal maps of the Northern Africa and Western Sahara scenes as observed by the two TIRS instruments at B10 and B11.

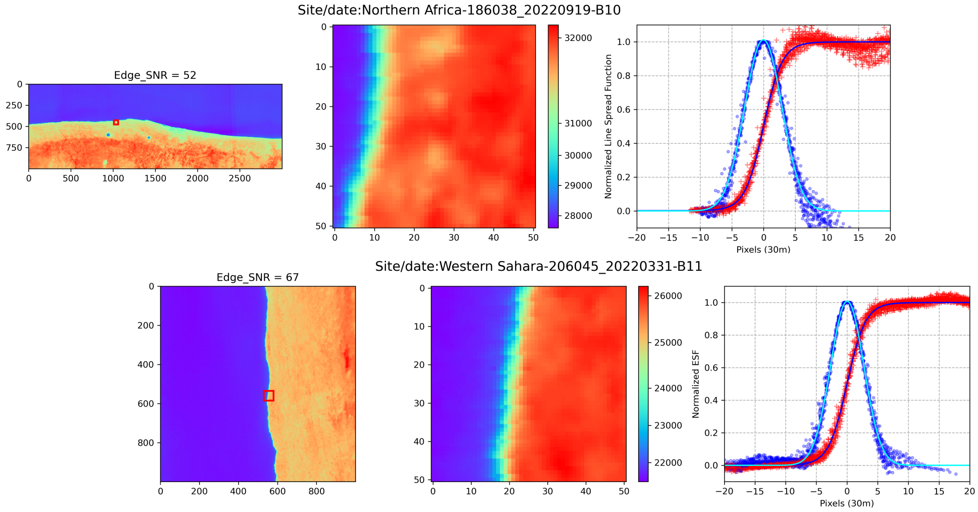

Figure 2. The top and bottom rows depict examples of the aligned edge profiles for the chosen sections of the edge for an along- and an across-track coastline (as well as their associated edge profiles in red, and their first-order derivative Line Spread Function in blue). The red square has been moved along the coastline to monitor the stability of the signal-to-noise and edge profiles on Landsat 8 TIRS images at the B10 band. The optimally found spot for metric measurements is chosen such that it has SNR$_{\text{edge}} > 50$ and a well-behaved aligned stack of profiles at the same spot for all of the qualified images of the scene throughout the year 2022. Once decided on the spot, it is affixed and used on all of the qualified images of Landsat 9 TIRS at B10 and B11.

The data used for this work are public and accessible through USGS's Earth Explorer data repository. The TIRS imagery in both Landsat missions are Tier 1 L1-TP, 30 m resampled (using the Cubic Convolution method described in the Landsat Level-1 processing details https://www.usgs.gov/landsat-missions/landsat-level-1-processing-details, accessed on 8 August 2024) TIF maps from the high quality (top-tier) Standard Terrain Corrected data products of collection 2. The data repository was queried over daytime nadir acquisitions by both TIRS instruments at both the B10 and B11 bands throughout 2022 (the first full-year overlap in their operations). Table 1 lists the chosen scenes and their properties. SAHA and LIBY are the across- and along-track scenes that are used for spatial metric trends. For these two scenes, the goal was to find the least cloudy image in each $\sim$30-day interval throughout 2022 acquired by both sensors, in order to assess the longer-term consistency of the measured spatial metrics. In the months throughout which no cloud-free images were acquired, the acceptable cloud coverage range was gradually increased up to $<$20% (and up to 40% in only two months for the Western Sahara site) until an image was found. Depending on the proximity of the cloud patch to our desired edge section, the chosen area for measurements would be moved by enough pixels away from the visible clouds to have a cloud-free section of the edge for our measurement. This was only necessary for one image in LIBY. The second set of along- and across-track sites (YMEN and OMAN) consisted of only 0% cloud coverage. Consequently, there are extended periods in 2022 during which no such image is available for one or both of these two sites. The idea behind choosing this limit for OMAN and YMEN was to have a set of "control" measurements in the along- and across-track directions, so that their metric fluctuations could not be speculated to have cloud-related reasons.

Before the final selection of the sites listed in Table 1, more coastlines in Western Australia, the western coast of Mexico, and the western coast of Greenland were examined. Those sites did not provide as high contrast and sharp boundaries as the chosen sites. Those sites also did not have as many <20% cloud covered scenes as SAHA and LIBY to make it possible to measure the trend of metrics during 2022. The numeric approach to find a perpendicular transect through the angled edge requires interpolation between crossing pixels. Combination of this numeric treatment with a lower contrast between the edge sides in some candidate scenes imposes a more selective approach in choosing the acceptable transects; a disadvantage that was more pronounced for the geometrically defined structures that are sharply angled with respect to the near-polar ground tracks of TIRS instruments (the Bay area bridges and Long Beach harbor were assessed before final selection). Our control scene, OMAN, happens to contain two geometrically defined structures in near-perfect perpendicular orientations with respect to each other and scan direction (see the Port of Duqm in Figure 3). The measured metrics from these structures were used to compliment the spatial performance assessment of the sensors in the absence of coastline edge imperfections. The selection of the dataset is further discussed in the context of the results in Section 4.

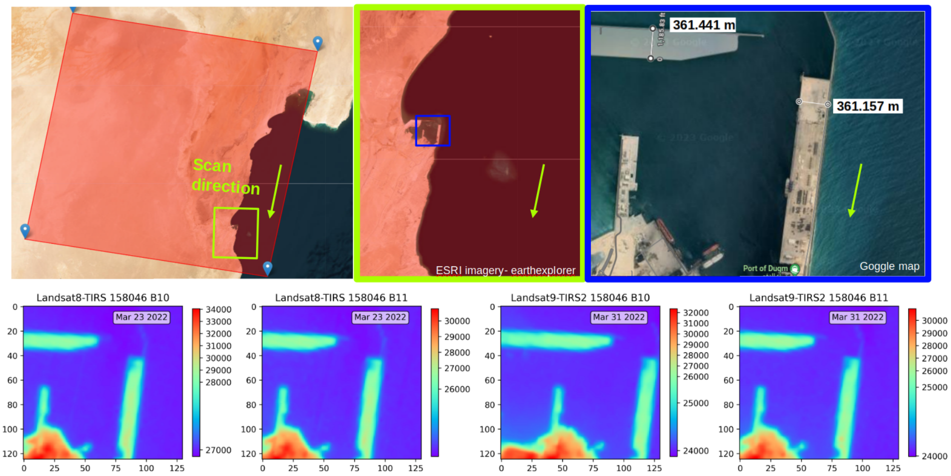

Figure 3. The layout of the Port of Duqm in the OMAN scene (Path/Row:158/46), with geometrically defined structures in near-perfect along- and across-track directions. The same spatial metrics that were measured in coastal edges were measured for these two structures on the same images, in which an edge in the coast Oman is measured (see the DQUM information in Tables 1 and 2). DUQM measurements were used as control data for along- and across-track metric trends in the coastal edges.

2.2. Measurement

The Edge Spread Function (ESF) is typically derived from analyzing the pixel values in each row of the chosen section of the image that includes the boundary between two relatively uniform yet significantly different sensor readings. Enough pixels on either side of the edge results in higher signal-to-noise edge profile (see Equation (1)). ESF describes how the imaging system spreads the intensity of a sharp edge in its scanned image. Mathematical representation of the distribution of intensities along the individual transects crossing the edge can quantify this spread. As the slanted edge with respect to the focal plane axes is scanned by the sensor, the intensity values are measured at various positions along the transect that passes through the edge. The slight tilt of the edge in the image is a key feature in this analysis. Transects passing through a very straight edge with respect to the image frame would all have similar intensity distribution profile, which results in an undersampled edge profile created from the individual transect profiles, while a tilted edge causes different transects to pass through the edge center at different spots, resulting in an oversampled edge profile with a heightened ability to record the true spatial characteristics of the imaging sensor from that image.

Some conventionally measured spatial characterization metrics that are derived from ESF include: (i) An edge slope that is measured at 40–60% of the normalized ESF. Higher edge slopes mean sharper edges. (ii) a Relative Edge Response (RER), which is a measure of the steepness of the edge profile, and is the measured as the drop in ESF amplitude when moving within ±0.5 pixel of the center of the edge. RER reveals how an imaging system responds to a change in contrast over one pixel. A higher RER indicates a steeper edge and sharper image. Blurry images have higher RERs. (iii) Line Spread Function (LSF) and its Full Width at Half Maximum (FWHM), which signifies the width of the ESF at half of its maximum value. LSF is the first order derivative of the the ESF, and represents the extent of the edge spread. Smaller FWHMs indicate sharper edges and better spatial quality. (iv) The Edge Extent that is defined as horizontal distance between 5% and 95% of the edge response (lower and upper edge points), and measures the contrast transition distance on either side of an edge. And (v) the Modulation Transfer Function (MTF), which is the Fourier Transformation of the LSF. The spatial frequency response evaluation method is typically developed in accordance with the International Standard ISO 12233:2000. MTF characterizes the imaging system's ability to reproduce different spatial frequencies or details in the image. Since MTF is a relative measure, it quantifies the ratio of the output modulation (i.e., contrast) to the input modulation for each spatial frequency. Higher MTF values at sampling frequency indicate a better ability to reproduce fine details and preserve contrast, while a lower MTF values indicates a reduction in detail reproduction and contrast. To quantify the shape of the MTF curves, three MTF shape parameters for each curve are measured: "MTF50", which is the spatial frequency at which the MTF drops to 50% of its maximum value; "MTF@1/2Nyq.", which is MTF at half Nyquist sampling; and "MTF@Nyq.", which represents MTF at the Nyquist frequency.

The next step after finding one qualified image per month for both TIRS instruments is to identify the optimal section of the edge for the metric measurement. $\mathrm{SNR}_{\mathrm{Edge}}$ was calculated for every 50 × 50 pixel array along the full stretch of a slightly tilted coastline with respect to the image frame (re square in Figure 2). This was performed for all of the qualified images of Landsat 8 TIRS at B10. The goal was to identify a section of the edge that consistently yields $\mathrm{SNR}_{\mathrm{Edge}} > 50$ throughout the year. Initially, a threshold of $\mathrm{SNR}_{\mathrm{Edge}} = 50$ was chosen, as it has been reported to be a relatively effective signal level [9]. Through numerous measurements of this factor from images captured under wide range of observing conditions and dates/times, this threshold was found to be sufficiently high to flag low-quality imagery, and not overwhelming high to unreasonably disqualify a reasonably sharp contrast. Equation (1) shows the definition of $\mathrm{SNR}_{\mathrm{Edge}}$ for the purpose of this study.

$$\mathrm{SNR}_{\mathrm{Edge}} = \frac{\mathrm{DN}_{\mathrm{bright}} - \mathrm{DN}_{\mathrm{dark}}}{(\mathrm{STD}[\mathrm{bright}] + \mathrm{STD}[\mathrm{dark}])/2} \tag{1}$$

After the most consistently high-quality section of the coastline in the scene was identified, images of the same section of the coast in all of the TIRS images of the coast throughout 2022 at both B10 and B11 were gathered. There were only a few cases where the optimum measurement section needed to be moved by a few pixels to bypass a thermal anomaly (i.e., either a cluster of abnormal pixel values due to moisture in the land or in the atmosphere over the area) in the thermal map of the scene. Two examples of this optimally chosen section of the coastline (indicated by a red square) are shown in the top- and bottom-left-most panels of Figure 2).

As described earlier, each row of the chosen 50 by 50 pixel array creates an edge profile. Those extracted profiles are then fit with a mathematical representation to locate the exact location of the edge with sub-pixel accuracy. A modified Fermi function ($F_m(x)$) following the behavior of the linear portion of the edge profile where the edge center is located:

$$F_m(x) = d + \frac{(b-d)}{1 + exp(-s(x-e))} + g(x-e). \tag{2}$$

where "d" represents the mean value on the dark side of the edge, "b" is the mean value on the bright side, "s" represents the slope of the linear portion of the profile (since the conventional definition of the edge slope as a spatial metric is limited to a smaller section around the edge center, the best-fit parameter "s" is not reported as the edge slope) and "e" the edge center location. The shifted profiles by the best-fit value of the edge center, e, create an oversampled edge profile that is then used to measure the spatial metrics defined earlier. An example of the aligned edge profiles in an along- and cross-track edges are shown in red in the top and bottom third panels of Figure 2. The aligned stack of normalized edge spread responses are then resampled using linear interpolation to a uniform pixel spacing of 0.1–0.01 for slope determination of its linear section (finer spacing than 0.1 did not improve the fit, nor key parameter determination). The slope is determined from the 40 to 60 percent section (0.4–0.6) of the normalized edge response. The same procedure is applied for the Edge Extent determination (defined as 10 to 90 percent of the normalized response). To remove biases from our results, measurements from aliased or blurred images were removed before the performance evaluation. A modified form of an image classifier, Q, to flag low-quality images was also measured in for the edge image (limited to the section where the edge profile is extracted). Q essentially connects the width of representative diffraction-limited PSF of an imaging sensor's produced image to its characteristic sampling distance. Q's modified definition, $\text{Q}_{\text{effective}}$, includes the detector, optics, electronics, motion effects, etc. that contribute to the shape of the LSF [19,20]:

$$\text{Q}_{\text{effective}} = \frac{\text{FWHM of the LSF}}{\text{GSD}_{\text{sensor}}}. \tag{3}$$

Using this definition, images with $\text{Q}_{\text{effective}} > 2$ are considered blurry, and those with $\text{Q}_{\text{effective}} < 1$ are considered aliased. A GSD_{sensor} of 100 m is applied for thermal imagery of Landsat 8 and 9. Prior to averaging all of the measurements of a particular spatial metric for each band and each sensor, any measurements associated with aliased and blurry images were removed from the set (see Section 3). Motivated by multiple published works on Edge Method analyses that had found SNR_{Edge} alone to be helpful, but not enough, for weeding out the low quality images, the temporal variation of all of the measured spatial metrics (edge slope, Edge Extent, half edge response, FWHM of LSF, and MTF shape parameters) were examined together with the trend of their associated Q and SNR_{Edge} (early examples of these trends were presented in [21]). For brevity, Figure 3 only shows the edge slope trend with these tow factors per image in both B10 and B11 bands. There are clearly edges with a high SNR_{Edge} that have yielded a low edge slope (i.e., less sharp edge,) and the Q factor of that edge categorized the image as blurry, which justifies the low edge slope. Enough such cases led us to look into a combination of $\text{SNR}_{\text{Edge}} > 50$ and $1 < \text{Q} < 2$ in our vetting process for qualified images to used for metric averaging (presented in Figures 8–10).

2.3. Uncertainties

A definition of the uncertainties of the monitored spatial metrics can differ between the responsible teams in charge of verification and validation of the mission requirements. But, conceptually, the main contributors to the observed uncertainties are largely agreed upon: (i) optical system mirrors and lenses introduce aberrations, diffractions, and distortions that affect the shape of the LSF and MTF; (ii) atmospheric turbulence, scattering, and absorption can also impact the sharpness and spatial response of the imaging system by affecting the edge profile and the derived LSF and MTF; (iii) uncertainties in the radiometric calibration process can propagate into the calibrated digital numbers (DN) in the final map and affect the derived parameters, including the edge slope and width; (iv) inherent errors from the processing algorithms for radiometric and geometric corrections and resampling steps also propagate into spatial response characteristics; and (v) the characteristics of the imaged scene, such as the presence of complex textures, high-frequency details, or areas with low contrast, can also influence the measured edge slope and width. Spectrally variable or complex scenes and, in the case of thermal images, scenes that include thermally variable/unstable areas, challenge accurate edge detection, and contribute to uncertainties in the LSF and MTF measurements.

In our approach adopted here, the metrics for each of the individual transects that pass through the edge (rows or columns) are calculated. The standard deviation of the calculated metrics from individual transects is reported as the uncertainty of the measured metric. The metric itself is calculated from the full stack of the aligned transects.

3. Results

Table 2 lists the average values of all of the measured metrics for each of the TIRS instruments, and for each scene in B10 and B11. After the metrics introduced in Section 2.2 were measured for each of the qualified images of each scene at each band, the trends of the metrics over the monitoring window (1 January 2022–1 January 2023) were assessed. Given that the two Landsat platforms are eight days out of phase, and to reduce observing condition differences between the two sensors, the closest acquisition dates for the qualified thermal images of a scene was chosen. The spatial metrics were measured on the same sections of the scenes, unless it was not possible due to existence of a patch of cloud in that particular spot, which prompted a few pixels to move to an adjacent section.

Table 2. Average metric values over the determined time window and number of scenes (see Table 1 and Section 2.1 for utilized data, and Section 2.2 for the definitions of the metrics). The Port of Duqm in the OMAN scene (Figure 3) was chosen for its near-perfect orientation of two geometrically defined structures in the along- and across-track directions (∼6° (XT), ∼3° (AT)). The measurements for the geometrically defined structures in DQUM, which only participated in the final stage of performance assessment as a potential comparison for the upper level of MTF amplitudes, had the edge imperfection of the shorelines removed.

Site	Edge Slope [1/pixels]	Rel. Edge Res. [pixels]	Half Edge Res. [pixels]	FWHM_{LSF} [pixels]	$\text{MTF}_{\text{Nyq.}}$
			Landsat 9 TIRS		
SAHA (B10)	0.608 ± 0.028	0.17 ± 0.013	8.616 ± 0.398	6.476 ± 0.159	0.098 ± 0.01
B11	0.669 ± 0.047	0.18 ± 0.018	8.472 ± 0.583	6.398 ± 0.280	0.112 ± 0.018
LIBY (B10)	0.604 ± 0.054	0.18 ± 0.013	7.985 ± 0.568	6.141 ± 0.208	0.138 ± 0.016
B11	0.612 ± 0.051	0.188 ± 0.015	7.788 ± 0.481	5.978 ± 0.147	0.153 ± 0.013
OMAN (B10)	0.721 ± 0.173	0.217 ± 0.047	7.222 ± 0.976	6.019 ± 0.401	0.171 ± 0.041
B11	0.671 ± 0.091	0.174 ± 0.034	8.50 ± 0.998	6.486 ± 0.538	0.123 ± 0.036
YEMN (B10)	0.664 ± 0.071	0.190 ± 0.031	7.648 ± 0.685	6.175 ± 0.293	0.146 ± 0.025
B11	0.624 ± 0.062	0.189 ± 0.029	7.756 ± 0.641	6.208 ± 0.277	0.140 ± 0.025
	$\text{XT} - \text{FWHM}_{\text{LSF}}$	$\text{XT} - \text{MTF}_{\text{Nyq.}}$		$\text{AT} - \text{FWHM}_{\text{LSF}}$	$\text{AT} - \text{MTF}_{\text{Nyq.}}$
DUQM (B10)	5.942 ± 0.275	0.306 ± 0.041		6.188 ± 0.054	0.227 ± 0.018
B11	5.80 ± 0.139	0.276 ± 0.020		5.658 ± 0.246	0.273 ± 0.043
			Landsat 8 TIRS		
SAHA (B10)	0.578 ± 0.026	0.185 ± 0.011	8.142 ± 0.273	6.167 ± 0.106	0.120 ± 0.008
B11	0.752 ± 0.0420	0.190 ± 0.018	7.921 ± 0.42	6.106 ± 0.158	0.127 ± 0.013
LIBY (B10)	0.572 ± 0.043	0.175 ± 0.012	7.518 ± 0.551	6.154 ± 0.173	0.128 ± 0.011
B11	0.564 ± 0.059	0.173 ± 0.021	9.033 ± 0.804	6.323 ± 0.439	0.1625 ± 0.043
OMAN (B10)	0.559 ± 0.112	0.177 ± 0.034	7.878 ± 0.685	6.205 ± 0.457	0.112 ± 0.034
B11	0.585 ± 0.104	0.189 ± 0.029	7.756 ± 0.641	6.241 ± 0.408	0.117 ± 0.038
YMEN (B10)	0.493 ± 0.049	0.149 ± 0.025	9.279 ± 0.641	6.848 ± 0.396	0.079 ± 0.021
B11	0.613 ± 0.072	0.180 ± 0.029	8.436 ± 0.859	6.476 ± 0.379	0.180 ± 0.029
	$\text{XT} - \text{FWHM}_{\text{LSF}}$	$\text{XT} - \text{MTF}_{\text{Nyq.}}$		$\text{AT} - \text{FWHM}_{\text{LSF}}$	$\text{AT} - \text{MTF}_{\text{Nyq.}}$
DUQM (B10)	5.813 ± 0.093	0.267 ± 0.017		5.618 ± 0.083	0.185 ± 0.025
B11	5.821 ± 0.212	0.295 ± 0.048		6.111 ± 0.148	0.211 ± 0.019

Keeping track of the SNR_{Edge} and $Q_{\text{effective}}$ values of each scene helped to verify the existence or strength of correlations between these two image quality factors and the measured spatial metrics. Besides learning that an image with $\text{SNR}_{\text{Edge}} < 50$ and $Q_{\text{effective}} > 2|Q_{\text{effective}} < 1$ would very likely not yield a tight edge profile and a small uncertainty envelope for its derived metrics, it was also observed that $\text{SNR}_{\text{Edge}} > 50$ does not strongly correlate with the best derived edge profiles from a scene (Figure 4 presents the variations of the edge slope, and the impact of considering SNR and Q together to correctly identify all of the images with less sharp edges).

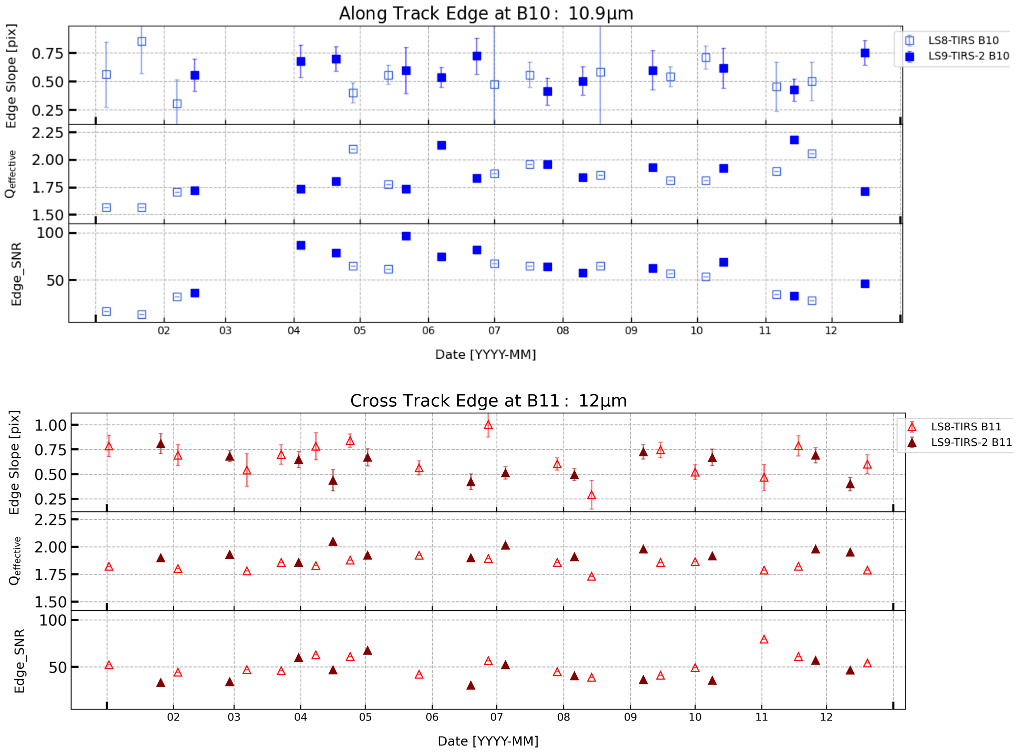

Figure 4. Trends of the spatial resolution metrics during 2022 for along-track scene (Northern Africa; 186/38): edge slope, FWHM of the LSF, and three MTF shape parameters (MTF value at the Nyquist frequency, the spatial frequency at which MTF drops to 50 percent, and MTF value at half Nyquist frequency). The measurements are color-coded for the two bands. Measurements associated with $\text{SNR}_{\text{Edge}} < 50$ or $Q_{\text{effective}} > 2$ are excluded from the averaged metrics in Figure 8.

Along with three MTF shape parameters (MTF50, MTF@1/2Nyq., and MTF@Nyq.), the changes to the full MTF curves with edge image orientation (across- versus along-track), edge contrast (SNR_{Edge}), spectral band and, ultimately, between the two thermal sensors were also investigated. Figure 5 illustrates an example of the stack of MTF curves for one scene, at B10 and at B11 of the Landsat 9 TIRS instrument for an along-track scene (Northern Africa; Path/Row: 186/38) during 2022.

The high SNR sections of TIRS images at B10 determine the focus areas for TIRS-2 images, with the possibility of some moderate pixel shifts when the thermal fluctuations of either side of the edge cause failure for the edge profile extraction algorithm. The gray curves in both panels are the measured MTF for images with $\text{SNR}_{\text{Edge}} > 50$. The MTF curves shown in colored dashed lines are for images with $\text{SNR}_{\text{Edge}} < 50$. Low SNR images do not participate in the averaging process for each of the defined metrics in Section 2.2.

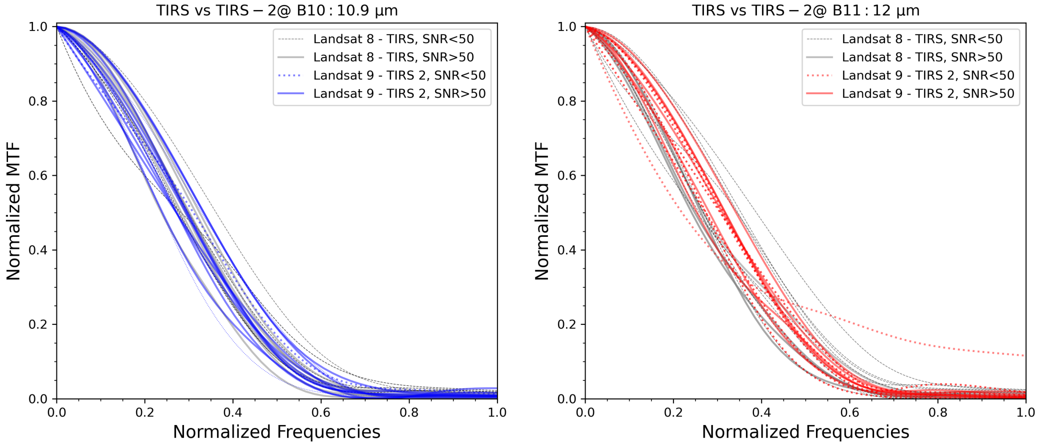

Figure 5. An example illustration of the comparison between the 12 normalized modulation transfer functions (MTFs) for the same section of the coastal edge in an along-track scene during 12 months of 2022 (the same comparison is performed for all of the along- and cross-track scenes presented in this work). In both panels, the black and gray curves show the MTFs from the Landsat 8 TIRS images, and the colored curves are the MTFs of the Landsat 9 TIRS-2 images.

Unlike the cross-track Western Sahara scene, finding a section of the image in the along-track scene (LIBY) that would yield $\mathrm{SNR}_{\mathrm{Edge}} > 50$ was not possible for every image. Looking at the gray and colored stacks of the MTF curves (representing TIRS and TIRS-2 imaging quality at an along-track scene), and then comparing them with the two stacks of MTFs from a cross-track scene, it appears that the MTF of along-track scenes are more sensitive to the SNR level than those of the across-track scene. Further, MTF curves of the images at similar SNR levels are more stable and consistent in B10 than B11. This observation remains the same for both the Landsat 8 and 9 TIRS.

Choosing Representative Metrics

It should be noted that all of the introduced metrics for the Edge Method in Section 2.2 stem from the shape of the edge profile. Consequently, the uncertainty of the measured edge slope impacts the uncertainties of other metrics, in accord with their definitions. Looking at our results, among the measured metrics, the uncertainty envelope of the edge slope affects the Edge Extent and Relative Edge Response the most, rendering them least sensitive to the slight performance differences between the two TIRS instruments. Therefore, those metrics were dropped, and the focus was shifted to investigating the trend of edge slope and FWHM of its derivative (LSF) to determine the more reliable of the two, with minimal failure rates to mimic the trend of the MTF shape change.

In Figure 6, the top two panels show the variations of the edge slope for the along-track scene of Northern Africa in B10 (the dashed and solid blue lines and shaded blue envelopes for TIRS and TIRS-2, respectively). The dashed and solid red lines and shaded yellow envelopes show the same metric for B11 in both sensors. The bottom two panels of this figure show the edge slope for the across-track scene in the Western Sahara. The overlaid open and filled diamonds and circles in all of the panels are the edge slopes for the control scenes in the along- and across-track directions. The trends of the edge slope in the top two panels (along-track scene) fluctuate over a wider range of values, and have comparatively higher uncertainties than their counterparts in the bottom two panels for the across-track scene (this can be seen in the relative position of the open markers with dashed lines for the Landsat 8 TIRS and the filled markers and solid lines for the Landsat 9 TIRS).

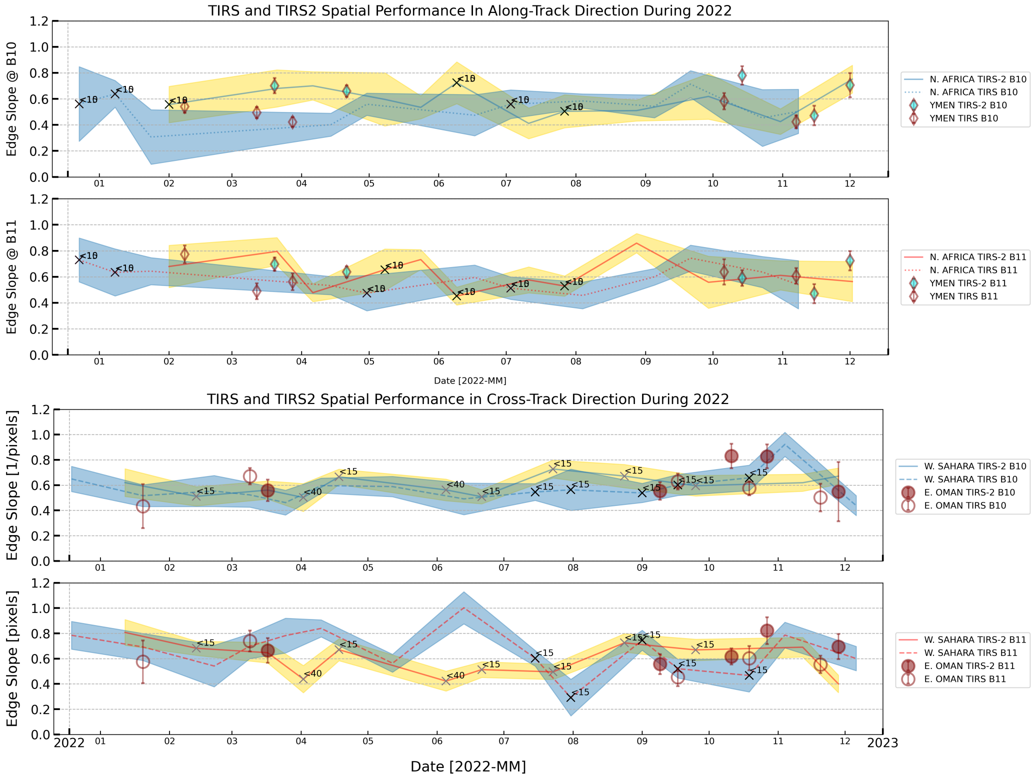

Figure 6. Top two panels: variations of the edge slope throughout 2022 for TIRS and TIRS-2 images of the along-track scene of Northern Africa at B10 (dashed and solid blue lines and shaded blue envelopes for TIRS and TIRS-2, respectively) and B11 (dashed and solid red lines and shaded yellow envelopes for TIRS and TIRS-2, respectively). Bottom two panels: same description as the top panels for the across-track scene in the Western Sahara.

The panels of Figure 7 follow the same color code and marker code as those in Figure 6. The trends of FWHM of the LSF in the top two panels, as compared to the two bottom panels, show more stability of this metric in across-track direction in both the Landsat 8 and Landsat 9 TIRS instruments.

The annotated "×" markers in Figures 6 and 7 flag the images with >0 cloud coverage. Our measurements for the presented data and for the test scenes prior to the final scene selection suggest that the impact of cloud coverage on spatial metrics is highly location-dependent, and seems to negatively impact the metrics on a case-by-case basis. Even though only the images for which the cloud coverage was calculated and reported to be <20% were used, a few images with a reported 0% cloud coverage were encountered that had a surprisingly low SNR_{edge} and widely spread edge profiles. No seasonal trends for metrics could be confidently derived from our data. The presence of seasonal dust particles and other atmospheric content pollutants that could have a moderate effect on the image clarity in the thermal bands (compared to more significant effects at visible and near IR) was considered. Difficulty in matching the image acquisition dates with the sparse reporting of the Aerosol content, and an often significant distance between the closest monitoring station and our chosen site, prevented further investigation of the possible impact.

Together, Figures 6 and 7 show that, in across-track scenes, TIRS instruments on both platforms perform similarly, with noticeable but insignificant spatial quality improvement in Landsat 9 TIRS (the solid lines and filled symbols in blue and red in all panels follow each other more closely than the dashed lines and open symbols in blue and red that represent the Landsat 8 TIRS metric values).

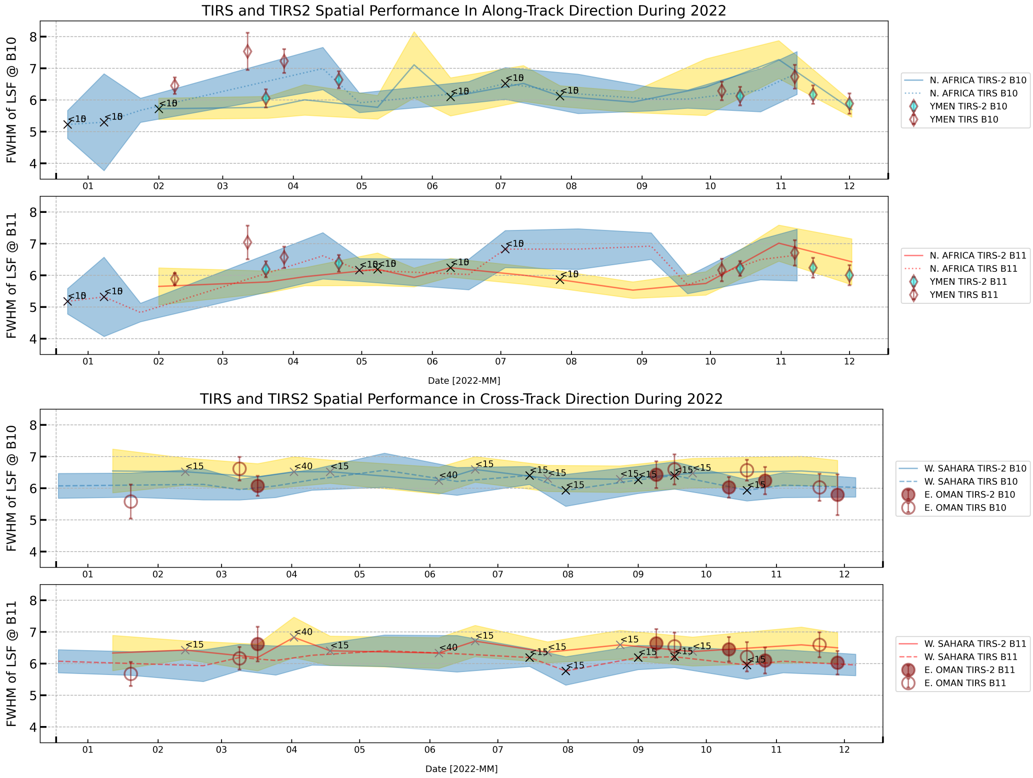

Figure 7. A similar style of graph to Figure 6 for the FWHM of the line spread function for the along- and cross-track scenes and their respective control scenes. See Sections 3 and 5 for more details and interpretations.

Lastly, comparing the edge slope and FWHM of the LSF for a representative metric, the trends of the FWHM of the LSFs for the across- and along-track scenes in both TIRS instruments appear more consistent with the associated metrics in their control scenes, less sensitive to edge spread profile imperfections and, therefore, are more reliable for long-term monitoring.

4. Discussion

Figure 8 encapsulates the results by visualizing the averaged metrics that are listed in Table 2. The first panel from the top left shows that the across-track edge slopes (open and filled triangles in blue and red) span a wider range of values than the along-track edge slopes (open and filled blue and red squares). Focusing on the colors, the edge slopes in B10 show more consistency over time than the edge slopes in the B11 images. The second panel from the top left shows that the FWHMs of the LSFs for across-track edges span over a slightly wider range of values, but their larger uncertainties have alleviated the statistical significance of this argument. Landsat 8 TIRS shows marginally tighter edge profiles (smaller FWHM) in the across-track direction, and Landsat 9 TIRS shows slightly sharper edges in the along-track direction.

The last two panels from the top left and the first panel from the bottom left show that the MTF shape parameters for the across-track scenes appear relatively consistent between the two TIRS instruments. Focusing on the open symbols in both red and blue colors, the along- and across-track scenes are more distinguishable in their MTF shape in the Landsat 8 TIRS. The last two panels from the bottom left show that the Edge Extent and Relative Edge Response are closely related to edge slope (a smaller edge slope means a wider Edge Extent and smaller Edge Response), but their wide uncertainty envelopes resulted in inconclusive patterns for both metrics in the across- and along-track directions (the open and filled triangles in red and blue, together, do not show an upward or downward trend). Figures 9 and 10 compare the average FWHM and MTFs for the natural coastlines with the same metrics for the two defined edges oriented in the along- and across-track directions, located in the Port of Duqm in the OMAN scene. The FWHM of the Line Spread Function (and its Fourier space equivalent, MTF) exhibited better consistency and stability over the monitoring period compared to the other measured spatial metrics (in Figure 4) and, therefore, were chosen to represent the overall quality of the naturally formed edges against the same averaged metric from a constructed edge.

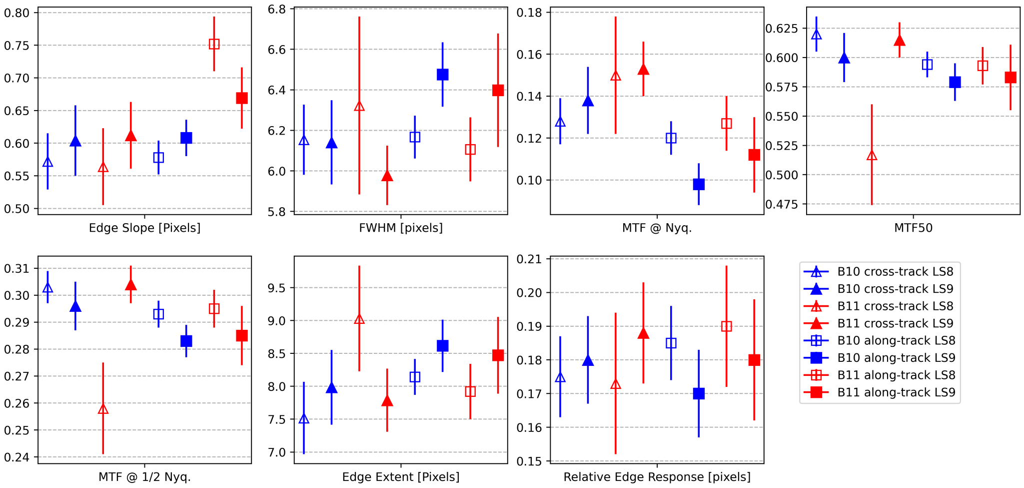

Figure 8. A summary diagram for the measured metrics, with the goal of illustrating the reason behind choosing a subset of the most effective metrics for better spatial performance comparisons. Each panel shows the average of one of the metrics per band and per site type (see Section 2.2 for their definition and the legend for colors and symbols). From top left to bottom right: Edge slopes in B10 (filled and open blue squares and triangles) show more consistency between edges in the along- and cross-track directions in both sensors than in B11 (filled and open red squares and triangles). The FWHM of LSF shows that the Landsat 8 TIRS shows marginally tighter edge profiles (smaller FWHM) in the across-track direction, and the Landsat 9 TIRS shows slightly sharper edges in the along-track direction. The MTF shape parameters (MTF@1/2Nyq., MTF50, and MTF@Nyq.) for across-track scenes appear relatively consistent between the two TIRS instruments. The large uncertainties in Edge Extent and Relative Edge Response have made them the least suitable metrics for detecting subtle changes with the spectral band and edge image orientation. The scenes with $\mathrm{SNR}_{\mathrm{Edge}} < 50$ and $Q_{\mathrm{effective}} > 2$ were removed prior to the averaging process.

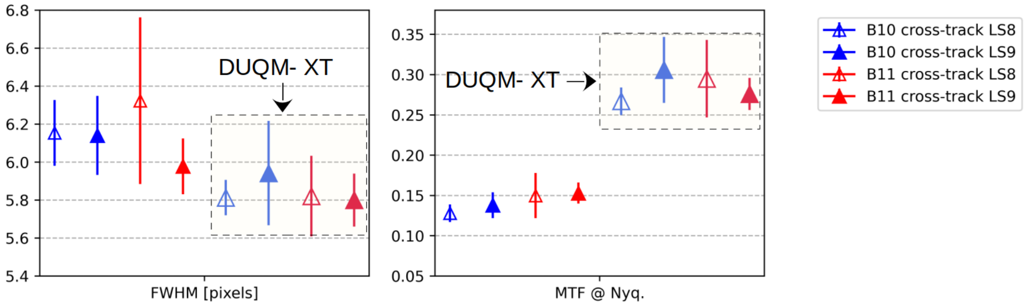

Figure 9. Average FWHM (**left**) and average MTF at Nyquist (**right**) for the monitored across-track scene throughout 2022 (values associated with blurry and low SNR scenes were excluded from the averaging process). Blue and red colors represent the two spectral bands, B10 and B11, respectively. The data points enclosed in the dashed square boxes in each panel are the same averaged metrics for the geometrically defined vertical structure in the Port of Duqm in the south-east of the OMAN scene (see Figure 3).

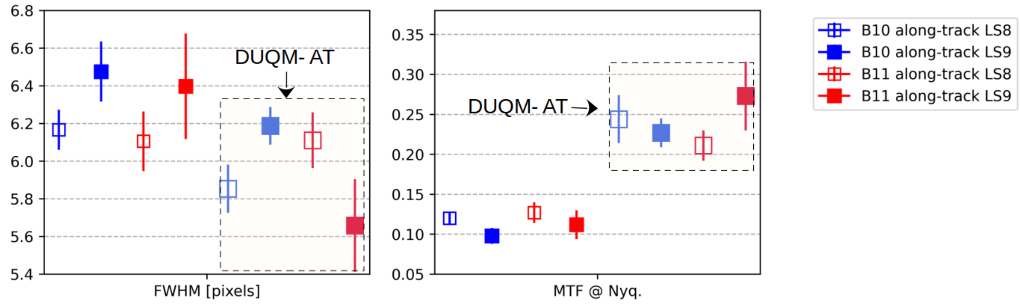

Figure 10. A similar figure to Figure 9, but for our monitored along-track scene throughout 2022. The data points enclosed in the dashed square boxes in each panel are the same averaged metrics for the horizontal structure in the Port of Duqm in the south-east of the OMAN scene (see Figure 3).

Figure 9 implies that both the FWHM and MTF show no sensitivity with spectral bands, and that the averaged across-track MTF at Nyquist for the constructed edge in the Port of Duqm is significantly higher than that of the coastlines. The FWHM of the Line Spread Function (and its Fourier space equivalent, MTF) exhibited better consistency and stability over the monitoring period, compared to other measured spatial metrics (Figure 4) and, therefore, were chosen to represent the overall quality of the naturally formed edges against the same averaged metric from a constructed edge. Both the FWHM and MTF show no variation when moving from B10 to B11. The averaged across-track MTF at Nyquist for a constructed edge in the Port of Duqm is noticeably higher than that of the coastlines.

The right panel of Figure 10 shows that the coastline edges in the Landsat 9 TIRS images for both bands spread wider than similar edges in the Landsat 8 TIRS by less than one pixel. The left panel shows that this slight difference in the averaged FWHM does not affect the averaged along-track MTF at Nyquist between the bands and between the instruments for coastline edges, but the bands and instruments were more distinguishable for the DUQM edges. The average MTF of the DUQM edge is marginally higher in the Landsat 9 TIRS images. The averaged across-track and along-track MTF at Nyquist for the constructed edge in the Port of Duqm is noticeably higher than those of the coastlines. The along-track MTFs appear to generally fall below the across-track MTFs, regardless of the spectral bands and edge type.

4.1. A Note on the Averaging Logic

Our key results in the last three figures of this paper are presented in the form of averaged metrics. Arguments can be made for and against using averaged metrics to validate the resolving power of sensors, or to compare the spatial performance of two sensors over time. One approach in reporting the best spatial performance of a sensor could come from the single closest measured metric to that of the system requirement, and independent from whether the metric is derived from a similar section of the scene, or the best section of the scene instead of a consistent location in all images of the same scene. Working with the average of a metric over an extended period of time, and remaining in almost similar section of the edge across all of the images for consistency of the edge shape, may create a wider range of measured values for a metric, as compared to when the best section of each image that yield the most similar metric value to the system requirement is chosen independently from other images. One may argue that this selective and case-by-case approach to metric measurements has a higher potential to report the true power of the instrument. However, this work upholds the repeatability and consistency of the measured metric over the “closest value to expected” for a consistent performance.

4.2. Data Volume and Diversity

In various trend analyses for performance assessment, increasing the data volume by either relaxing the qualifications or expanding the monitoring period has a proven record of enhancing meaningful trends or dissolving biased or nonphysical patterns. A quantitative assessment of the trends of the edge slope and FWHM of the LSF for a sample of cloud-free images of two along- and cross-track scenes from the Landsat 8 TIRS over its first decade of operation (2013–2023) triggered a sensitivity assessment of the spatial metrics in thermal bands. Neither trend showed conclusive evidence of seasonal fluctuations. As detailed in Section 2.1, more shorelines around the globe were examined than those listed in Table 1. The listed shorelines were found to have a more consistently high contrast than others (such as coasts of Mexico or Australia). Measurements from more natural edges that had fewer qualified images throughout the year did not provide new insights into the spatial performance of the two sensors. In other words, these two instruments are too similar, and these metrics at these wavelengths, measured for such imperfect edge targets, are too insensitive to subtle performance drifts, to be able to more accurately compare the spatial quality of their data products. The lack of access to reference data during the sensor calibration process, and to the native resolution imaging products (which are not made publicly available), and having no sufficiently detailed knowledge of the properties of the imaging system (in order to select and optimize the parameters of the MTF compensation algorithm), prevented us from using the MTF compensation algorithm to enhance the MTF signal. An increased MTF signal could provide the needed statistical significance for subtle differences between the ability of each sensor to reproduce fine imaging details stored at closely separated spatial frequencies. Or, alternatively, the ability to more confidently vote on the similarity of the two metrics.

5. Conclusions

Understanding the spatial quality of a remote sensor is one component of successfully interpreting the operational on-orbit sensor imagery. The spatial performance is typically determined at pre-launch in controlled environments, using carefully selected and well-understood edge targets in the optical test equipment. Once a sensor is in orbit, it is challenging to assess and monitor its spatial performance. One major challenge is in identifying Earth surface features with a sharp thermal transition that are imaged regularly, oriented advantageously to the sensor orbit, and have uniformity in response on either side of the edge. Coastlines in the desert regions of northern Africa and the Arabian Peninsula are found to be useful for spatial performance analysis, as they have suitable thermal contrasts, extended geometrical lengths in a variety of orientations, and low cloud-contamination images can be collected during most months throughout the year. Methodologies were developed to identify and select edge targets for TIRS images useful for both across- and along-track spatial analysis. A combination of the $Q_{\mathrm{effective}}$ and SNR_{Edge} factors are used in this work to quantify the quality of the thermal contrast around the edge, and we found them to be effective for the exclusion of low-quality images that would increase the uncertainty of the spatial metrics, yet they appeared insufficient to guarantee high spatial quality.

On-orbit analysis of the selected desert coastal scenes using the same methodology over one year of their shared operational period demonstrated that the Landsat 8 TIRS and Landsat 9 TIRS spatial performance are consistent and nearly identical in both the across- and along-track directions. The nearly identical, but 8-day out-of-phase, orbits for the two platforms allowed for the same set of coastal edge targets to be used for both sensors. The TIRS instruments used both the edge slope and Edge Extent to define spatial performance requirements at pre-launch. The similarity of their performance is more pronounced in across-track directions when looking at the trends of edge slopes and the LSF's FWHM for both sensors over the monitoring period. The trend of averaged MTF shape parameters imply that Landsat 8 TIRS distinguishes between along- and across-track scans more than the Landsat 9 TIRS, but this level of difference does not contradict the reports on Landsat 8 TIRS consistency with its spatial quality requirement. Given the analyzed imaging data for this study, our metric evaluation showed that the line spread function and its Fourier counterpart sufficiently carry the information about the condition of the edge. The averaged MTF values at Nyquist for thermally stable, geometrically defined edges in built structures (such as harbor platforms and wide bridges) showed that along-track MTFs appear to generally fall below across-track MTFs, regardless of the spectral band and edge type (an effect largely associated with and explained by motion blur that is present along the sensor track and not across track). However, throughout the course of MTF measurements from the derived Line Spread Function, the vulnerability of the MTF shape parameters to the LSF imperfections became evident. The noticeable (but not significant) differences between the along- and across-track can be attributed to the challenges of the on-orbit use of natural, but imperfect, edge targets. The edge slope, Edge Extent, and MTF are sensitive to the the opto-mechanical changes to the instruments in the controlled lab environment with thermally uniform targets in parameterized scan sequences. However, measuring those same metrics to verify nominal spatial performance of the now on-orbit instrument can most likely result in poorer estimates of its true performance. This has been the reason that this study opts for stability of metrics over time, in place of emphasis on similarity with pre-launch metric values. It can be operationally rewarding to plan ahead for the calibration validation sites of the future thermal sensors, with calibration validation sites that include geometrically defined, thermally stable boundaries in the along and across orientations with respect to their scanning track. Using harbor structures often requires compromises on usable edges and extended areas on either side of the thermal boundaries, as well as numeric treatment of the crossing transects through the edge to the appropriately perpendicular direction.

Author Contributions: S.E.K. led the calculations and reporting of the measurement. B.N.W. and K.J.T. contributed and oversaw the final write up and the rest of the team members (M.Y., D.J.L., M.H.T. and N.V.) were involved in auditing the calculations and its progress toward the compilation of the final version of the manuscript in regular meetings and helped crafting the conclusions to this study. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The data that support the findings of this study are openly available in Landsat dataset repository at https://earthexplorer.usgs.gov/, accessed on 8 August 2024.

Acknowledgments: S.E.K. is grateful for the illuminating discussions with Matthew Montanaro and Raviv Levy on a number of technical details related to Landsat TIRS instruments and their associated data treatment.

Conflicts of Interest: The authors S. Eftekharzadeh Kay, B. N. Wenny, M. Yarahmadi and M. H.Tahersima were employed by the company Science Systems & Applications, Inc. The remaining authors declare that the research was conducted without commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Reuter, D.C.; Richardson, C.M.; Pellerano, F.A.; Irons, J.R.; Allen, R.G.; Anderson, M.; Jhabvala, M.D.; Lunsford, A.W.; Montanaro, M.; Smith, R.L.; et al. The Thermal Infrared Sensor (TIRS) on Landsat 8: Design Overview and Pre-Launch Characterization. *Remote Sens.* **2015**, *7*, 1135–1153. [CrossRef]
2. Jhabvala, M.; Reuter, D.; Choi, K.; Jhabvala, C.; Sundaram, M. QWIP-based thermal infrared sensor for the Landsat Data Continuity Mission. *Infrared Phys. Technol.* **2009**, *52*, 424–429. [CrossRef]
3. Hair, J.H.; Reuter, D.C.; Tonn, S.L.; McCorkel, J.; Simon, A.A.; Djam, M.; Alexander, D.; Ballou, K.; Barclay, R.; Coulter, P.; et al. Landsat 9 Thermal Infrared Sensor 2 Architecture and Design. In Proceedings of the IGARSS 2018—2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018 ; pp. 8841–8844. [CrossRef]
4. Montanaro, M.; McCorkel, J.; Tveekrem, J.; Stauder, J.; Mentzell, E.; Lunsford, A.; Hair, J.; Reuter, D. Landsat 9 Thermal Infrared Sensor 2 (TIRS-2) Stray Light Mitigation and Assessment. *IEEE Trans. Geosci. Remote Sens.* **2022**, *60*, 1–8. [CrossRef]
5. Eon, R.; Wenny, B.N.; Poole, E.; Eftekharzadeh Kay, S.; Montanaro, M.; Gerace, A.; Thome, K.J. Landsat 9 Thermal Infrared Sensor-2 (TIRS-2) Pre- and Post-Launch Spatial Response Performance. *Remote Sens.* **2024**, *16*, 1065. [CrossRef]
6. Eon, R.; Gerace, A.; Falcon, L.; Poole, E.; Kleynhans, T.; Raqueño, N.; Bauch, T. Validation of Landsat-9 and Landsat-8 Surface Temperature and Reflectance during the Underfly Event. *Remote Sens.* **2023**, *15*, 3370. [CrossRef]
7. Pearlman, A.; McCorkel, J.; Montanaro, M.; Efremova, B.; Wenny, B.; Lunsford, A.; Simon, A.; Hair, J.; Reuter, D. Landsat 9 Thermal Infrared Sensor 2 pre-launch characterization: Initial imaging and spectral performance results. In Proceedings of the Earth Observing Systems XXIII, San Diego, CA, USA, 19–23 August 2018; Volume 10764, p. 1076406. [CrossRef]

8. Gerace, A.; Montanaro, M. Derivation and validation of the stray light correction algorithm for the thermal infrared sensor onboard Landsat 8. *Remote Sens. Environ.* **2017**, *191*, 246–257. [CrossRef]
9. Wenny, B.N.; Helder, D.; Hong, J.; Leigh, L.; Thome, K.J.; Reuter, D. Pre- and Post-Launch Spatial Quality of the Landsat 8 Thermal Infrared Sensor. *Remote Sens.* **2015**, *7*, 1962–1980. [CrossRef]
10. Barsi, J.A.; Schott, J.R.; Hook, S.J.; Raqueno, N.G.; Markham, B.L.; Radocinski, R.G. Landsat-8 Thermal Infrared Sensor (TIRS) Vicarious Radiometric Calibration. *Remote Sens.* **2014**, *6*, 11607–11626. [CrossRef]
11. Montanaro, M.; Levy, R.; Markham, B. On-Orbit Radiometric Performance of the Landsat 8 Thermal Infrared Sensor. *Remote Sens.* **2014**, *6*, 11753–11769. [CrossRef]
12. Forshaw, M.R.B.; Haskell, A.; Miller, P.F.; Stanley, D.J.; Townshend, J.R.G. Spatial resolution of remotely sensed imagery A review paper. *Int. J. Remote Sens.* **1983**, *4*, 497–520. [CrossRef]
13. Carnahan, W.H.; Zhou, G. Fourier transform techniques for the evaluation of the Thematic Mapper line spread function. *ISPRS J. Photogramm. Remote Sens.* **1986**, *52*, 639–648. [CrossRef]
14. Plass, W.; Maestle, R.; Wittig, K.; Voss, A.; Giesen, A. High-resolution knife-edge laser beam profiling. *Opt. Commun.* **1997**, *134*, 21–24. [CrossRef]
15. Blonski, S.; Pagnutti, M.A.; Ryan, R.; Zanoni, V. In-flight edge response measurements for high-spatial-resolution remote sensing systems. In Proceedings of the Earth Observing Systems VII, Seattle, WA, USA, 7–11 July 2002; Volume 4814, pp. 317–326. [CrossRef]
16. Thome, K.; Lunsford, A.; Montanaro, M.; Reuter, D.; Smith, R.; Tesfaye, Z.; Wenny, B. Calibration plan for the Thermal Infrared Sensor on the Landsat Data Continuity Mission. In Proceedings of the Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XVII, Orlando, FL, USA, 25–29 April 2011; Volume 8048, p. 804813. [CrossRef]
17. Viallefont-Robinet, F.; Helder, D.; Fraisse, R.; Newbury, A.; van den Bergh, F.; Lee, D.; Saunier, S. Comparison of MTF measurements using edge method: Towards reference data set. *Opt. Express* **2018**, *26*, 33625–33648. [CrossRef]
18. Masaoka, K. Edge-based modulation transfer function measurement method using a variable oversampling ratio. *Opt. Express* **2021**, *29*, 37628–37638. [CrossRef] [PubMed]
19. Pagnutti, M.; Blonski, S.; Cramer, M.; Helder, D.; Holekamp, K.; Honkavaara, E.; Ryan, R. Targets, methods, and sites for assessing the in-flight spatial resolution of electro-optical data products. *Can. J. Remote Sens.* **2010**, *36*, 583–601. [CrossRef]
20. Alonso, K.; Bachmann, M.; Burch, K.; Carmona, E.; Cerra, D.; de los Reyes, R.; Dietrich, D.; Heiden, U.; Hölderlin, A.; Ickes, J.; et al. Data Products, Quality and Validation of the DLR Earth Sensing Imaging Spectrometer (DESIS). *Sensors* **2019**, *19*, 4471. [CrossRef] [PubMed]
21. Eftekharzadeh, S.; Wenny, B.N.; Thome, K.J. On-orbit spatial calibration performance of Landsat 8 and 9 thermal infrared sensors. In Proceedings of the Infrared Remote Sensing and Instrumentation XXXI, San Diego, CA, USA, 20–25 August 2023; Volume 12686, p. 1268603. [CrossRef]