Article

Detection of Change Points in Pseudo-Invariant Calibration Sites Time Series Using Multi-Sensor Satellite Imagery

Neha Khadka [1,2], **Cibele Teixeira Pinto** [2,*] **and Larry Leigh** [2]

1   Department of Electrical Engineering and Computer Science, South Dakota State University, Brookings, SD 57007, USA; neha.khadka@jacks.sdstate.edu
2   Imaging Center, Image Processing Laboratory, South Dakota State University (SDSU), Brookings, SD 57007, USA; Larry.Leigh@sdstate.edu
*   Correspondence: Cibele.TeixeiraPinto@sdstate.edu; Tel.: +1-605-688-5217

Abstract: The remote sensing community has extensively used Pseudo-Invariant Calibration Sites (PICS) to monitor the long-term in-flight radiometric calibration of Earth-observing satellites. The use of the PICS has an underlying assumption that these sites are invariant over time. However, the site's temporal stability has not been assured in the past. This work evaluates the temporal stability of PICS by not only detecting the trend but also locating significant shifts (change points) lying behind the time series. A single time series was formed using the virtual constellation approach in which multiple sensors data were combined for each site to achieve denser temporal coverage and overcome the limitation of dependence related to a specific sensor. The sensors used for this work were selected based on radiometric calibration uncertainty and availability of the data: operational land imager (Landsat-8), enhanced thematic mapper (Landsat-7), moderate resolution imaging spectroradiometer (Terra and Aqua), and multispectral instrument (Sentinel-2A). An inverse variance weighting method was applied to the Top-of-Atmosphere (TOA) reflectance time series to reveal the underlying trend. The sequential Mann–Kendall test was employed upon the weighted TOA reflectance time-series recorded over 20 years to detect abrupt changes for six reflective bands. Statistically significant trends and abrupt changes have been detected for all sites, but the magnitude of the trends (maximum of 0.215% change in TOA reflectance per year) suggest that these sites are not changing substantially over time. Hence, it can be stated that despite minor changes in all evaluated PICS, they can be used for radiometric calibration of optical remote sensing sensors. The new approach provides useful results by revealing underlying trends and providing a better understanding of PICS' stability.

Keywords: PICS; sequential Mann–Kendall (SQMK); Mann–Kendall (MK); virtual constellation (VC); inverse variance weighting; Landsat; MODIS; Sentinel

1. Introduction

Satellite remote sensing is used to acquire information related to the Earth's resources and to monitor changes that occur in the land surface, coastal and oceanic ecosystems. Radiometric calibration of remote sensing sensors is the foundation for using its data in a further quantitative application. The complete radiometric calibration of a sensor onboard a satellite typically consists of both pre-launch and post-launch radiometric calibration. Pre-launch radiometric calibration involves evaluating sensor and onboard calibrators in a controlled environment under a variety of operating conditions [1]. These onboard calibrators and sensors are characterized and calibrated using the Système International (SI) traceable standards to achieve the needed accuracy [2]. Nevertheless, once the satellites are launched into space, exposure to ultraviolet radiation and other electrical, mechanical, and thermal effects, may degrade their calibration over time [3]. In consequence, continuous re-calibration is required during the entire sensor lifetime (post-launch calibration).

Post-launch radiometric calibration of remote sensing sensors can be achieved using artificial and natural ground targets on the Earth surface. This technique is known as

Vicarious Calibration, in which the term "vicarious" refers to calibration methods that are conducted without the use of systems that are directly mounted onboard the satellite. Pseudo-Invariant Calibration Site (PICS) based calibration is one of the most widely used vicarious calibration methods. PICS are regions on the Earth that exhibit minimal change over long periods and can be used in three forms of radiometric calibration: (a) long term trending of the sensor radiometric gains; (b) cross-calibration; and (c) more recently, absolute calibration [4–6].

For instance, Morstad [4] in one of his previous research used PICS to observe the accurate trending of satellite instruments. The results indicated the sites' consistency and usefulness for long-term trending of the Landsat series, both historically and in the future. The Committee on Earth Observation Satellites (CEOS) referenced PICS has also been used to evaluate the improvement in the long-term on-orbit radiometric calibration stability of the Terra Moderate Resolution Imaging Spectroradiometer (MODIS) [7]. The results show excellent calibration stability (long-term drift within 2%) of the MODIS sensor. Vuppula [8] used Landsat-8 data from six Saharan desert sites, normalized to the Libya-4 reference, and merged the normalized data into a "Super PICS" dataset to achieve a high temporal resolution calibration dataset. The resulting uncertainties were within the satellite calibration range of 3% for all spectral bands. Previous studies have also emphasized the use of PICS for cross-calibration of Landsat-8 OLI and Landsat-7 ETM+ and analyses of this approach show the uncertainty for all analogous bands of the two sensors agree within ±2% [6]. Chander [9] performed cross-calibration of the Advanced Land Imager (ALI) with Landsat-7 and MODIS using near-simultaneous surface observations based on the image data from areas observed over four desert sites (Libya 4, Mauritania 2, Arabia 1, and Sudan 1). Results showed agreement of ALI with MODIS and Landsat-7 within 5%. Besides, PICS has been used to assess the cross-calibration between Sentinel-2A and Landsat-8 of eight corresponding matching spectral bands with Landsat-8 as reference. The cross-calibration results show that seven corresponding bands' radiometric differences are consistent to Landsat-8 within 1% or better, except on the cirrus band [10]. Similarly, Govaerts [5] did his initial work developing an absolute calibration model using PICS with geostationary satellites. Results indicated 3% accuracy for the red and Near Infrared (NIR) band and 6% accuracy for other bands, which provided the motivation for the extension of these techniques. Helder [11] introduced the concept of an empirical absolute calibration model by employing one of the PICS site (Libya 4) as a test site, Terra MODIS as the calibrated radiometer, Hyperspectral Observations from Earth Observing (EO-1) Hyperion for the spectral signature and used Landsat-7 ETM+ image data for model validation. The model showed accuracy within 3% in the visible and 6% in the Short-Wave Infrared (SWIR) region, which suggested that absolute calibration using PICS, provided improved on-orbit calibration capabilities for satellite sensors. Later, this work was extended using five additional Saharan Desert PICS (Egypt 1, Niger 1, Niger 2, Sudan 1 and Libya 1) to generate the absolute calibration models for these sites with accuracy and precision comparable to or better than Libya 4 model [12].

The use of PICS sites for calibration has an underlying assumption that these sites are invariant over time. Consequently, any changes detected in lifetime trending of the radiometric gain are caused by the sensor response. However, the site's temporal stability has not been assured due to the lack of emphasis given to validate the foundational assumption. Recently, Tuli [13] developed an explicit assessment of PICS' temporal stability and demonstrated a long-term temporal variability in some of the PICS. However, the work only addressed the overall temporal stability of PICS without indicating when the behavior of the PICS data has changed in the time series. The motivation for this study is precisely related to the lack of work involving a finer assessment of PICS, specifically, associated with an evaluation of the presence of a change point in PICS time series. Change point detection is the method of finding changes in a dataset when a property of the time series changes. It is important to emphasize that several studies related to PICS trend analysis have been published, however, the possible time where the behavior of PICS data

has changed has not yet been evaluated. In this context, this work aims to expand the trend analysis of PICS by focusing on detecting significant shifts lying behind the PICS time series trend. The sites investigated in this work are PICS that are commonly used in calibration analysis at the South Dakota State University Image Processing Laboratory (SDSU IPLAB): Libya 1, Libya 4, Sudan 1, Egypt 1, Niger 1, and Niger 2. The goal of change point assessments is to describe the nature and degree of the known change and to identify a stable dataset over a time period and make its use for a successful radiometric calibration.

To monitor the long-term temporal stability of the PICS, virtual constellation approach was implemented. There are currently more than 150 Earth-observing satellites that are orbiting the Earth, carrying sensors that measure different sections of the visible, microwave, and infrared regions of the electromagnetic spectrum. The increase in the number of remote sensing sensors provides the community a unique opportunity to combine data sets. Combining data from two or more sensors into a single data set creates a "virtual constellation" [14]. According to CEOS, a virtual constellation is a "coordinated set of space or ground segment capabilities from different partners that focuses on observing a particular parameter or a set of parameters of the Earth system" [15]. The three significant advantages of combining data from multiple sensors into a single seamless time series are: (a) it increases the temporal resolution of the dataset; (b) it overcomes the dependence limitation of the potential trend that exists in any particular sensor; and (c) it allows a revealing new understanding of PICS. The multiple sensor data fusion involves reducing the effects of differences in the sensor's radiometric performance, such as spectral response and acquisition viewing/illumination geometry. The sensors applied in this work are Operational Land Imager (OLI) on-board Landsat-8 satellite, Enhanced Thematic Mapper (ETM+) on-board Landsat-7, Moderate Resolution Imaging Spectroradiometer (MODIS) on-board Terra, MODIS on-board Aqua, and Multispectral Instrument (MSI) on-board Sentinel-2A. These sensors are selected based on the radiometric calibration uncertainty and availability of the data.

The paper is structured as follows: Section 2 provides the work methodology along with the satellite sensors and the location of the six PICS utilized in this work. It also includes the pre-processing steps applied to the satellite sensors imagery to create a virtual constellation and the non-parametric statistical tests used for the analysis. Section 3 presents the results and the interpretation of the results for all six PICS. Section 4 presents the discussion and, finally, concluding remarks are presented in Section 5.

2. Materials and Methods

2.1. Satellite Sensor Overview

A brief overview of satellite sensors used for this work is described in this section. The image data used in this work came from five orbital platforms: Landsat-7, Landsat-8, Sentinel-2A, Aqua, and Terra. For implementing the virtual constellation approach, it is necessary to combine sensors with similar spectral characteristics to match them with minimal processing requirements [16]. Hence, the sensors from different orbital platforms used in this work have similar spectral coverage, enabling synergistic use. The common spectral bands for the five sensors are denominated blue, green, red, NIR, SWIR 1, and SWIR 2. These six spectral bands are listed in Table 1.

Table 1. Spectral bands of satellite sensors.

Sensor/Satellite	Bandwidth (nm)					
	Blue	**Green**	**Red**	**NIR**	**SWIR 1**	**SWIR 2**
OLI/Landsat-8	452–12	533–590	636–673	851–879	1567–1651	2107–2294
ETM+/Landsat-7	441–514	519–611	631–692	772–898	1547–1748	2064–2346
MSI/Sentinel-2A	470–524	504–602	649–680	855–875	1569–1658	2113–2286
MODIS/TERRA	459–479	545–565	620–670	841–876	1628–1652	2105–2155
MODIS/AQUA	459–479	545–565	620–670	841–876	1628–1652	2105–2155

2.1.1. Landsat ETM+/OLI

The Landsat series of sensors have acquired the longest continuous global imagery of the Earth's surface. Landsat-8 and Landsat-7 are part of the Landsat series and were launched on 11 February 2013, and 15 April 1999. The revisit period of the Operational Land Imager (OLI) onboard Landsat-8 is 16 days, which is eight days out of phase with the Enhanced Thematic Mapper plus (ETM+) onboard Landsat-7. OLI is a push broom imager that has nine spectral bands covering a 185 km swath width and ETM+ is a whiskbroom imager with 8 spectral bands and a 183 km swath width. Both OLI and ETM+ have a spatial resolution of 30 m in their multispectral bands. Landsat-7 has a local overpass time of 10:30 am, which is close to that of Landsat-8 (i.e., 10:13 am) [17]. The ETM+ has an estimated uncertainty of ±5%, and OLI has an estimated post-launch calibration uncertainty of 2% and pre-launch calibration uncertainty of approximately 3% in reflectance products [18,19].

2.1.2. Sentinel MSI

The Multi-Spectral Instrument (MSI) onboard Sentinel-2A is part of the European Space Agency's (ESA, Noordwijk, Netherland) Copernicus program that has been contributing Earth remote sensing data since 2015 [19]. It provides images of the entire Earth every 10 days with spatial resolution varying between 10 m to 60 m for 13 spectral bands. MSI is a push broom imager that has a 295km swath width. The orbit is Sun-synchronous at 786 km altitude with a 10:30 am equatorial crossing time [20]. The data obtained from the onboard Multi-Spectral Instrument (MSI) is comparable to other well-calibrated sensors such as the OLI. The OLI and MSI showed stable radiometric calibration of approximately 2.5% with consistency between matching spectral bands [21].

2.1.3. Terra and Aqua MODIS

The Moderate Resolution Imaging Spectroradiometer (MODIS) onboard Terra and Aqua satellites is one of the key instruments for the National Aeronautics and Space Administration (NASA, Washington DC, USA) Earth Observing System (EOS) providing data for more than 20 years [22]. The NASA MODIS Characterization Support Team (MCST) is responsible for on-orbit calibration, characterization, and sensor operation. Terra has been collecting imagery since 18 December 1999, and Aqua since 04 May 2002. The MODIS sensor is designed to image the Earth in 36 spectral bands with a spatial resolution ranging from 250 m to 1000 m. The instrument is a whiskbroom scanner with a wide field of view of 2330 km and a view zenith angles up to ±45°. Due to the regular assessment of the sensor's on-orbit performance, both Terra and Aqua MODIS are viewed as well-calibrated sensors. Terra and Aqua MODIS are also commonly called the AM and PM series of satellites because of their equatorial crossing time of 10:20 am and 1:30 pm. They can view the entire globe once in every two days [22]. The estimated radiometric calibration uncertainties of the MODIS TOA reflectance products are 2% [23,24].

2.2. Study Area

The sites investigated in this work are the Pseudo-Invariant Calibration Sites located in North Africa. These are the six SDSU IPLAB PICS (Libya 4, Niger 1, Sudan 1, Niger 2, Egypt 1, and Libya 1), that demonstrate better temporal stability and have larger areas as compared to other CEOS referenced sites [25]. The region identified by the PICS Normalization Process (PNP) algorithm developed by the SDSU IPLAB as "optimal" regions was used as the region of interest. This region exhibits temporal, spatial, and spectral variability of 3% or less [8]. Table 2 shows the corresponding latitude and longitude coordinates defining the Region of Interest (ROI) along with the Sentinel UTM/WGS84 projection tiles and Landsat Worldwide Reference System 2 (WRS-2) path and row.

Table 2. Location of study areas, including the WRS-2 Path/Row, UTM/WGS84 Tiles, and ROI coordinates.

PICS	WRS-2 Path/Row	MSI Tile	Minimum Latitude	Maximum Latitude	Minimum Longitude	Maximum Longitude
Libya 4	181/40	34RGS	28.38	28.81	23.09	23.86
Niger 1	189/46	32QNH	20.28	20.53	9.19	9.52
Sudan1	177/45	35QND	21.40	21.75	27.81	27.59
Niger 2	188/45	32QPJ	21.25	21.47	10.38	10.71
Egypt 1	179/41	35RMK	26.91	27.13	26.31	26.62
Libya 1	187/43	33RUH	24.55	24.86	13.32	13.66

2.3. Image Pre-Processing

Radiometrically and geometrically corrected standard MSI data product images referred to as Level 1C (L1C) are available at the Copernicus Open Access Hub (https://scihub.copernicus.eu/ (accessed on 2 October 2020)). Landsat-8 OLI and Landsat-7 ETM+ Collection-1 Level-1 data are available at the United States Geological Survey (USGS) Earth Explorer website (https://earthexplorer.usgs.gov (accessed on 27 March 2020)). For the MODIS data, Collection 6.1 image data products for both Terra and Aqua can be accessed from MODAPS Web Services (https://modwebsrv.modaps.eosdis.nasa.gov/ (accessed on 16 August 2020)). Imagery data for the corresponding Sentinel-2A MSI, Landsat-7 ETM+, Landsat-8 OLI, Terra MODIS, and Aqua MODIS for the selected PICS were downloaded and later retrieved from the existing SDSU IPLAB archive. The pre-processing flowchart of this work is shown in Figure 1.
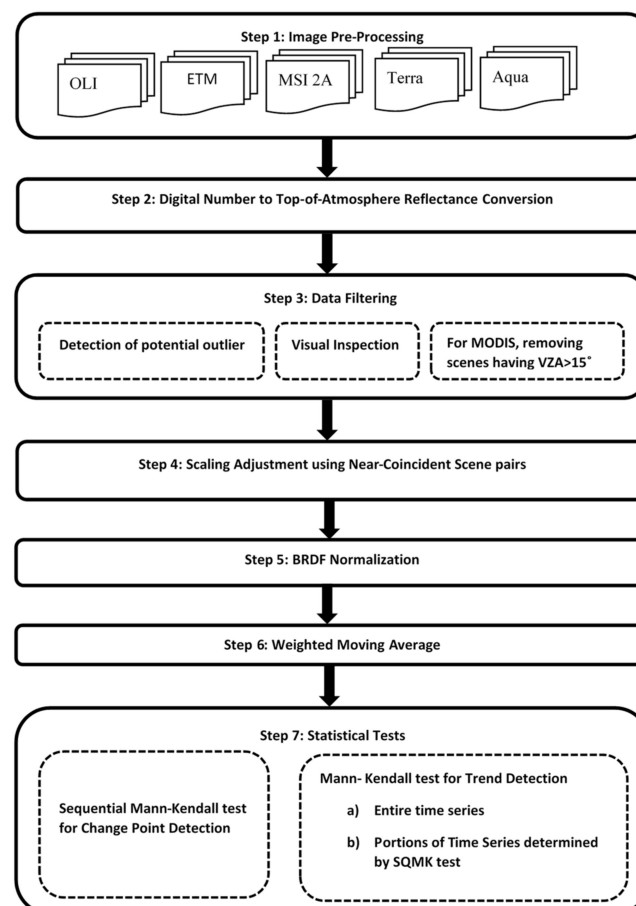

Figure 1. General flowchart for change point detection of PICS. VZA is the View Zenith Angle; BRDF is Bidirectional Reflectance Distribution Function; SQMK is Sequential Mann–Kendall.

2.3.1. Conversion to TOA Reflectance

Landsat Collections Level-1 data product consists of calibrated and quantized scaled Digital Numbers (DN) representing the multispectral image data. These reflective band DN's were converted into TOA reflectance using the rescaling coefficients from the MTL file [26,27].

$$\rho'_\lambda = M_\rho * Q_{cal} + A_\rho \tag{1}$$

where, ρ'_λ is TOA reflectance without the solar angle correction , M_ρ is band-specific multiplicative rescaling factor from the metadata, Q_{cal} is quantized and calibrated standard product pixel values DN, and A_ρ is band specific additive rescaling factor from the metadata. The solar angle correction of the TOA reflectance was calculated by using the formula:

$$\rho_\lambda = \frac{\rho'_\lambda}{\cos(SZA)} \tag{2}$$

where, ρ_λ is TOA reflectance without the solar angle correction, and SZA is solar zenith angle. The conversion of the Sentinel-2A MSI DN's to TOA reflectance was done by using a single constant for scaling, which accounts for the exo-atmospheric irradiance, Earth–sun distance, and the solar angle cosine correction:

$$\rho_\lambda = \frac{DN_{cal}}{g} \tag{3}$$

where, DN_{cal} is calibrated DN value (pixel value), and g = 1000 is the scale factor. The MODIS image data products were processed to TOA reflectance using the calibration parameters from the metadata for the same region of interest used for other sensors at their specified spatial resolution. The input calibration parameters' values were transformed into the reflectance scales and reflectance offsets by using Equation (4).

$$\rho \cos(\theta)_{B,T,FS} = reflectance\_scales_B(SI_{B,T,FS} - reflectance\_offsets_B) \tag{4}$$

where, $\rho \cos(\theta)_{B,T,FS}$ is reflectance product for reflected solar bands, $SI_{B,T,FS}$ is level 1B product scaled integer, and reflectance_scales$_B$ and reflectance_offsets$_B$ are parameters deduced from Level 1B calibration parameters.

2.3.2. Data Filtering to Remove Cloud, Cloud Shadow, and Large Sensor View Angles

After calculation of the cosine corrected mean TOA reflectance value for the Region of Interest (ROI) of each image, filtering was required to obtain cloud and cloud shadow-free imagery. A quality assessment band is included with Landsat Level-1 data products in which each pixel in the Quality Assessment band consists of a decimal value, which represents bit-packed combinations of sensor, atmosphere, and surface conditions. Thus, the quality assessment band data can be used to filter out the cloud-contaminated pixels from ETM+ and OLI imagery. The pixel value of 672 (clear pixels) of the quality assessment band was used to form a cloud binary mask for removing cloudy pixels from ETM+ imagery and a pixel value of 2720 for OLI imagery [3]. Likewise, for Sentinel, cloud mask data from a cloud mask file was used to remove the cloudy pixels for all the bands with different spatial resolutions. Some outliers remained in the dataset even after applying the cloud mask using the quality assessment band and the cloud mask data. Thereby, an empirical sigma (σ) filtering approach was used for identifying outliers. Basically, a scene was considered a potential outlier when the mean TOA reflectance value of the scene exceeds one sigma (± 1σ) threshold. Subsequently, a visual inspection was carried out for the images that were considered as a potential outlier. When cloud, cloud shadow, or other artifacts were identified during the visual inspection of the image within the ROI for any spectral band of an image, the entire scene was discarded from the analysis. Figure 2 illustrates a potential identified outlier and the image corresponding to the point highlighted by the red circle.

As can be seen, the red rectangle represents the ROI on the image, and it contains clouds, therefore this image was discarded from the analysis.
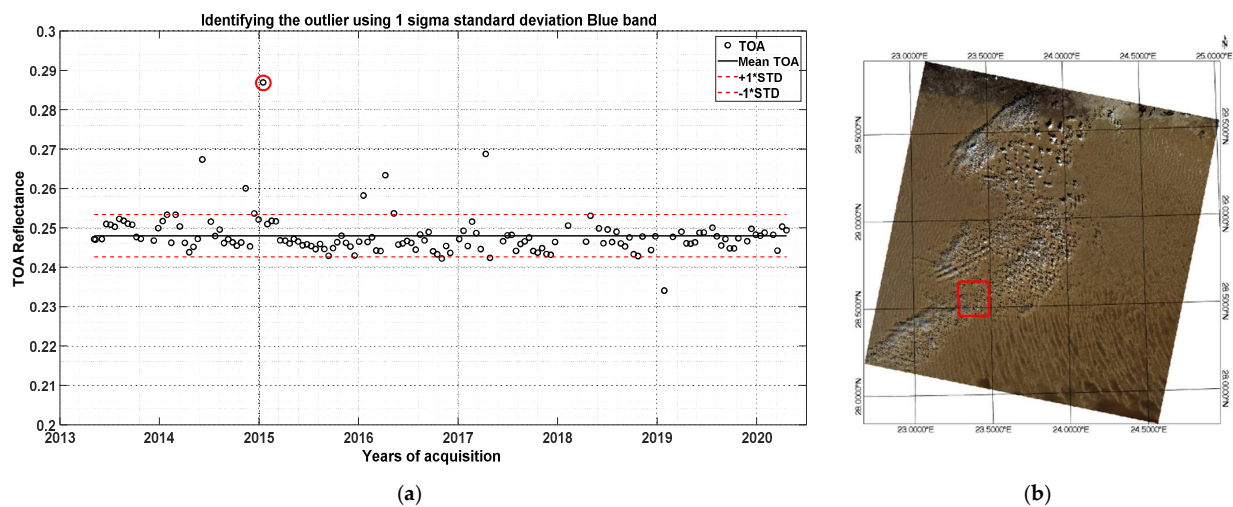

(a)　　　　　　　　　　　　　　　　　　　　　(b)

Figure 2. (**a**) Graph to illustrate the identification of outlier using 2-sigma standard deviation method; and (**b**) visual inspection of the image considered as a potential outlier.

All satellite sensors used in this work had view zenith angles of 7.5 degrees or less except MODIS. The Field of View (FOV) of the MODIS sensor is $\pm 49.5°$, which can increase the BRDF differences between the sensors due to uneven illumination and reflectance. Therefore, for MODIS, the scenes having a large view zenith angle (VZA > 15°) were excluded from the analysis in order to minimize the BRDF differences between the sensors.

2.3.3. Scaling Adjustment Factor (Near Coincident Scene Pairs)

The five sensors used in this work collect data in similar spectral regions; however, there are some differences in the Spectral Response Function (SRF) (see Table 1). These differences in SRF between sensors may generate a systematic error during the process of combining the data. Consequently, for mitigating the effects of differences in SRF between sensors, a procedure to achieve spectral consistency and to correct the absolute difference in the data was employed [13,16].

The TOA reflectance generated from ETM+, MODIS, and MSI sensors were scaled to match the OLI sensor by applying a Scaling Adjustment Factor (SAF), i.e., the TOA reflectance of each sensor was normalized by using the adjustment factor. The SAF was calculated by taking the mean ratio of cosine corrected TOA Reflectance values from near coincident acquisitions with the OLI (see Equation (5)). Near coincident acquisition refers to the scenes that are imaged within a maximum acceptable window of days. For MODIS and Landsat-7 ETM+, 'near-coincident' refers to the scene pairs that are imaged approximately eight days apart as they are eight days out of phase with the OLI. The scene pairs that are imaged within the window of four days were used for Sentinel-2A since they provide optimal scaling factor with minimum differences in between the sensors.

$$\text{SAF} = \frac{\rho_{\text{OLI}}}{\rho_{\text{ETM+/MSI/MODIS}}} \tag{5}$$

where, SAF is scaling adjustment factor, ρ_{OLI} is mean TOA reflectance from OLI sensor, and $\rho_{\text{ETM+/MSI/MODIS}}$ is mean TOA reflectance from ETM+, MSI, or MODIS sensor.

2.3.4. Bidirectional Reflectance Distribution Function (BRDF) Normalization

PICS are not perfect Lambertian surfaces, and consequently, directional reflectance effects are present in satellite remote sensors reflectance retrievals due to variability in the geometry of acquisition, especially in longer wavelength bands. The TOA reflectance

is dependent on the illumination and the viewing geometry, which vary significantly from one acquisition to another. To address the impacts due to the different solar and viewing geometries on the derived TOA reflectance, a Bidirectional Reflectance Distribution Function (BRDF) normalization was applied to the dataset.

Several approaches have been reported in the literature to model the BRDF data. Liu [28] employed a BRDF model based only on the solar zenith angle for cross-calibration between MODIS and MVIRS. Schlapfer [29] also proposed a normalization method for sensors having a wide field of view based on the solar zenith, view zenith, and relative azimuth angles (i.e., the difference between the solar and view azimuth angles). Likewise, Mishra [25] derived an empirical linear BRDF model based on solar and view zenith angles using near nadir images of Terra MODIS over Libya 4. Farhad [30] developed a linear BRDF model where the reflectance's were modeled as a quadratic function of view zenith angles. This model was further extended by Kaewmanee [31] with quadratic terms for all four angles (solar zenith/azimuth and view zenith/azimuth), including the interaction terms. This model provided the best BRDF model characterization and the best after normalization uncertainty. For this work, Kaewmanee's 15 coefficient quadratic model derived from four angles was chosen for BRDF normalization (see Equation (6)).

$$
\begin{aligned}
\rho_{\text{model}} = {} & \beta_0 + \beta_1 Y_1^2 + \beta_2 X_1^2 + \beta_3 Y_2^2 + \beta_4 X_2^2 + \beta_5 X_1 Y_1 + \beta_6 X_1 Y_2 + \beta_7 X_2 Y_2 + \beta_8 X_2 Y_1 + \beta_9 Y_1 Y_2 + \beta_{10} X_1 X_2 \\
& + \beta_{11} X_1 + \beta_{12} Y_1 + \beta_{13} X_2 + \beta_{14} Y_2
\end{aligned}
\tag{6}
$$

where, β_0, β_1, β_2, ... are the model coefficients, and Y_1, X_1, Y_2, X_2 are Cartesian coordinates representing the planar projections of the solar and sensor angles originally given in spherical coordinates:

$$Y_1 = \sin(\text{SZA}) \times \sin(\text{SAA}) \tag{7}$$

$$X_1 = \sin(\text{SZA}) \times \cos(\text{SAA}) \tag{8}$$

$$Y_2 = \sin(\text{VZA}) \times \sin(\text{VAA}) \tag{9}$$

$$X_2 = \sin(\text{VZA}) \times \cos(\text{VAA}) \tag{10}$$

where, SZA is solar zenith angle, SAA is solar azimuth angle, VZA is view zenith angle, and VAA is view azimuth angle. The BRDF-normalized TOA reflectance for each sensor was calculated using:

$$\rho_{\text{BRDF−normalized}} = \frac{\rho_{\text{obs}}}{\rho_{\text{model}}} \times \rho_{\text{ref}} \tag{11}$$

where, ρ_{obs} is observed mean TOA reflectance from each scene , ρ_{model} is model predicted TOA reflectance, and ρ_{ref} is TOA reflectance with respect to a set of "reference" solar and sensor position angles. For this analysis, the reference SZA, SAA, VZA, and VAA angles were chosen by taking the mode of the corresponding SZA, SAA, VZA, and VAA angles from all processed scenes. The reference angle chosen for six PICS for the normalization procedure is listed in Table 3.

Table 3. Reference angles for PICS.

SZA	SAA	VZA	VAA
38°	144°	3°	103°

In the course of selecting an appropriate TOA reflectance harmonization algorithm for the TOA reflectance dataset from five different sensors, the use of a single BRDF model worked well in the data by giving a minimum coefficient of variation (lesser amount of dispersion around the mean) in the normalized dataset. Although previous work [13] used different BRDF models for different sensors for the same site, it is well known that the directional effect is a property related to the target. Hence, it would be more reasonable to normalize the combined TOA reflectance data using a single BRDF model for a site with a common reference angle rather than using a different BRDF model for different sensors data for the same site.

2.4. Statistical Analysis

2.4.1. Inverse Variance Weighted Moving Average Method

Intrinsic in the TOA reflectance data collected over time is some form of random variation. Smoothing (or averaging) is a technique applied to time series to reduce the effect caused due to random variation in the data. The idea is to remove the noise from the data and expose the signal to the underlying trend, seasonal and cyclical components in the time series. There are several smoothing techniques, which involve averaging methods and exponential smoothing methods. Simple averaging and averaging by weights fall under averaging methods, where simple averaging implies taking the average of the observations and averaging by weights implies assigning equal or unequal weights to all the observations and then averaging them. In this work, weighted moving averages were applied, with the weights based on each observation's uncertainty. Non-parametric statistical analyses were applied to the resulting smoothed series [32].

The method of weighting the TOA reflectance data by uncertainties and applying the moving average is referred to as the inverse variance weighted average method. Calculating a moving average involves producing a new time series where the values comprise the weighted average of observations in the original time series. There is a need for weighting the data before applying statistical tests because there may be an over-representation of certain characteristics in the dataset. For instance, the test might be more biased towards the OLI sensor as we are using OLI as a reference for the scaling adjustment. On the other hand, having a larger number of observations than other sensors could also influence the test results. Weighting compensates for biases in the data. The moving average technique requires a specific window size, which defines the number of original observations used to calculate the moving average value.

In this work, a moving average with a window size of twelve months with seven days shifting was applied to reduce the seasonal effects, smooth out the short-term fluctuations in the time series data, and highlight long-term trends. For this application, the weights were allocated by making each point's weight inversely proportional to the square of its uncertainty [33] (see Equation (13)). Hence, data points in the time series with less uncertainty are given more weight than data points with higher uncertainty values (see Equation (12)).

$$\rho_{\text{weighted}} = \frac{\sum_{i=1}^{n} w_i \rho_i}{\sum_{i=1}^{n} w_i} \tag{12}$$

$$w_i = \frac{1}{\sigma_{\text{total}}^2} \tag{13}$$

where, ρ_i is virtual constellation TOA reflectance time series, w_i is weight, and σ is total uncertainty associated with each TOA reflectance. The total uncertainty associated with each observation of the time series was calculated taking into account four sources of uncertainties: (1) spatial uncertainty; (2) BRDF model uncertainty; (3) scaling adjustment factor uncertainty; and (4) sensor calibration uncertainty (see Equation (14)).

$$\sigma_{\text{total}} = \sqrt{\sigma_{\text{spatial}}^2 + \sigma_{\text{BRDF}}^2 + \sigma_{\text{SAF}}^2 + \sigma_{\text{sensor}}^2} \tag{14}$$

The spatial uncertainty was calculated considering the variation of mean TOA reflectance within the ROI. BRDF model uncertainty was evaluated by taking the residual errors between the observed and the predicted TOA reflectance. Scaling adjustment factor uncertainty is the ratio of the standard deviation to the mean scaling adjustment factor. Lastly, the sensor calibration uncertainties for different sensors used for this work are listed in Table 4.

Table 4. Sensor calibration uncertainty for five different sensors [21,22,25,27,28].

Sensors	Reflectance Product Uncertainty
L7- ETM+	5%
L8-OLI	2%
Sentinel 2A	2.5%
Terra MODIS	2%
Aqua MODIS	2%

2.4.2. Statistical Test for Change Point Detection (Sequential Mann–Kendall Test)

After applying the inverse variance weighting method, change point detection test is applied to the dataset. The purpose of the change point detection test is to investigate possible dates when the behavior of time series data has changed. The sequential version of the Mann–Kendall test [34] is a non-parametric test applied to time-series data to detect potential change points. This analysis has been increasingly employed in different fields, including finance, aerospace, meteorology, and climate changes. Bisai [35] examined the fluctuations of trend and the abrupt change point of atmospheric temperature time series in the Kolkata observatory using the sequential Mann–Kendall test. Kundu [36] also employed the Mann–Kendall test and Sen's slope estimator to determine the annual variability in the meteorological parameters of Western Rajasthan.

The test uses the ranked values of the original data in the analysis and locates the beginning year of the trend. It sets up two series, one progressive $u(t)$ and another retrograde $u'(t)$ by computing the rank values y_i from the original values $(x_1, x_2, x_3 \ldots, x_n)$ and the magnitudes of y_i $(i = 1,2,3 \ldots ,n)$ are compared with $y_j (j = 1,2,3 \ldots ,i\text{-}1)$. For each comparison, if $y_i > y_j$, they are counted and denoted by n_i. The test statistic t_i is given by:

$$t_i = \sum_{j=1}^{i} n_i \tag{15}$$

The distribution of test statistic t_i has a mean as

$$E(t_i) = \frac{i(i-1)}{4} \tag{16}$$

and variance as

$$Var(t_i) = \frac{i(i-1)(2i+5)}{72} \tag{17}$$

The sequential values for the series were calculated for each of the test statistic variables t_i using:

$$u(t_i) = \frac{t_i - E(t_i)}{\sqrt{var(t_i)}} \tag{18}$$

Here, the progressive sequential statistic $u(t_i)$ is estimated using the original time series $(x_1,x_2,x_3 \ldots, x_n)$ whereas the values of a retrograde sequential statistic $u'(t_i)$ are estimated in the same way but starting from the end of the series $(x_n,x_{n-1}, x_{n-2} \ldots, x_1)$. The standardized variable $u(t)$ has a zero mean and unit standard deviation and fluctuates around zero [37]. When the values of $u(t_i)$ (progressive series) and $u'(t_i)$ (retrograde series) are plotted as a function of time, the intersection point of these curves represents the approximate turning point of trend within the time series. When either the progressive or retrograde curve exceeds a certain threshold before or after the intersection point, this trend turning point (change point) is considered significant at the corresponding level [35]. The threshold value in this study is ±2.58 (99% significance level), with the crossing point estimating the year at which the trend begins.

2.4.3. Statistical Test for Trend Analysis (Mann–Kendall Test)

The Mann–Kendall test is a non-parametric trend detection technique that tests the null hypothesis H_0: data are independently and randomly ordered so that there is no monotonic trend present against the alternative hypothesis H_a: there is a monotonic trend in the data [38,39]. In other words, the Mann–Kendall test analyzes data collected over time for consistently monotonic trends (increasing or decreasing trends).

For this work, the Mann–Kendall test was applied to the same time series using two different strategies: (a) by dividing the time series into different regions based on the change points detected by the sequential Mann–Kendall test and applying the test in those regions to identify the slope magnitude of the trend; and (b) applying the test to the entire time series to look into the presence of an overall trend in every band. The mathematical theory behind the Mann–Kendall test is that it compares the difference in signs between earlier and later data points. The concept is that if a trend is present, the sign values will tend to increase or decrease based on the direction of the trend. Let $x_1, x_2, x_3 \ldots, x_n$ represent n data points where x_j represents the data point at time j. The Mann–Kendall test statistic is given by:

$$S = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \operatorname{sgn}(x_j - x_i) \tag{19}$$

where:

$$\operatorname{sgn}(x_j - x_i) = \begin{cases} 1 \text{ if } x_j - x_i > 0 \\ 0 \text{ if } x_j - x_i = 0 \\ -1 \text{ if } x_j - x_i < 0 \end{cases} \tag{20}$$

The initial value of the Mann–Kendall statistic, S, is assumed to be 0, and each data value is compared to all the subsequent data values in the ordered time series. If the data value from a later time period is higher than the data value of an earlier time period, then S is incremented by 1. If the data value from the later time period is lower than the data value of the earlier time period, S is decremented by 1. A very high positive value of S indicates an increasing trend, and a very low negative value indicates a decreasing trend [40]. This test was performed at a 99% significance level.

3. Results

The results obtained at each step of the methodology for Sudan 1 are presented in Section 3.1 and the summary of results for rest of the sites are presented in Section 3.2.

3.1. Stability Analysis of Sudan 1

3.1.1. Scaling Adjustment Factor

To implement the virtual constellation approach, TOA reflectance data was homogenized using the scaling adjustment factor as described in Section 2.3.3. Figure 3 illustrates TOA reflectance from the combined satellite sensors over time for Sudan 1 site without (Figure 3a) and with (Figure 3b) scaling adjustment factor. The offsets seen in Figure 3a caused by atmospheric effects, differences in the SRF, and the spectral signature of the ground target, of the individual sensors, are reduced after applying the scaling adjustment factor.
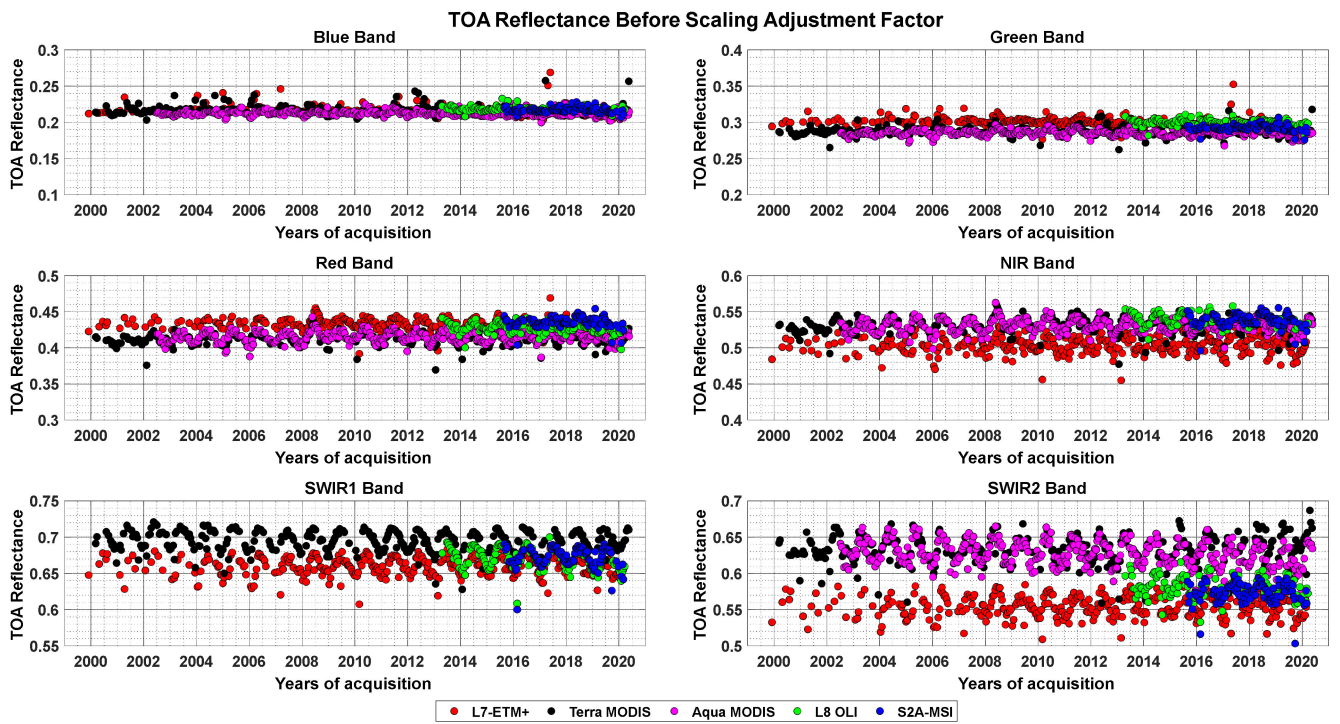
(**a**)
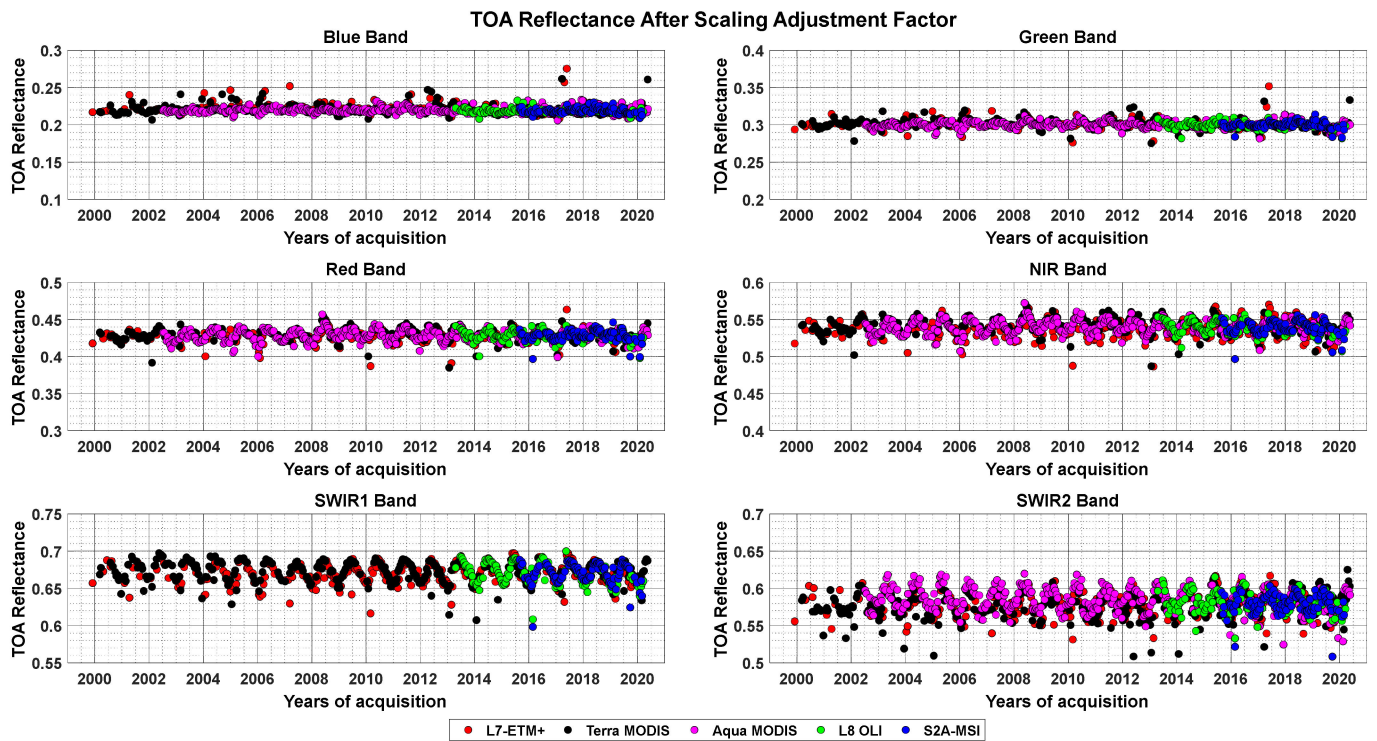
(**b**)

Figure 3. Temporal trend of before (**a**) and after (**b**) scaling adjusted TOA reflectance over Sudan 1 site for the combined dataset (ETM+, Terra MODIS, Aqua MODIS, OLI, and MSI).

3.1.2. Bidirectional Reflectance Distribution Function (BRDF) Normalization

The adjusted TOA reflectance from the combined sensor data was BRDF normalized in order to minimize seasonal effects. Figure 4 shows the TOA reflectance predicted by the 15-coefficient BRDF model explained in Section 2.3.4. The residual error of this model was estimated by taking the mean percentage difference between the observed value and value predicted by the BRDF model for all the sites, which is presented in Table 5. The residual errors are expected to be close to zero when the model is working well for the data. From the table, it is evident that the residuals are close to zero except for the blue and SWIR 2 band. This might be due to the atmospheric and water absorption features of these bands.
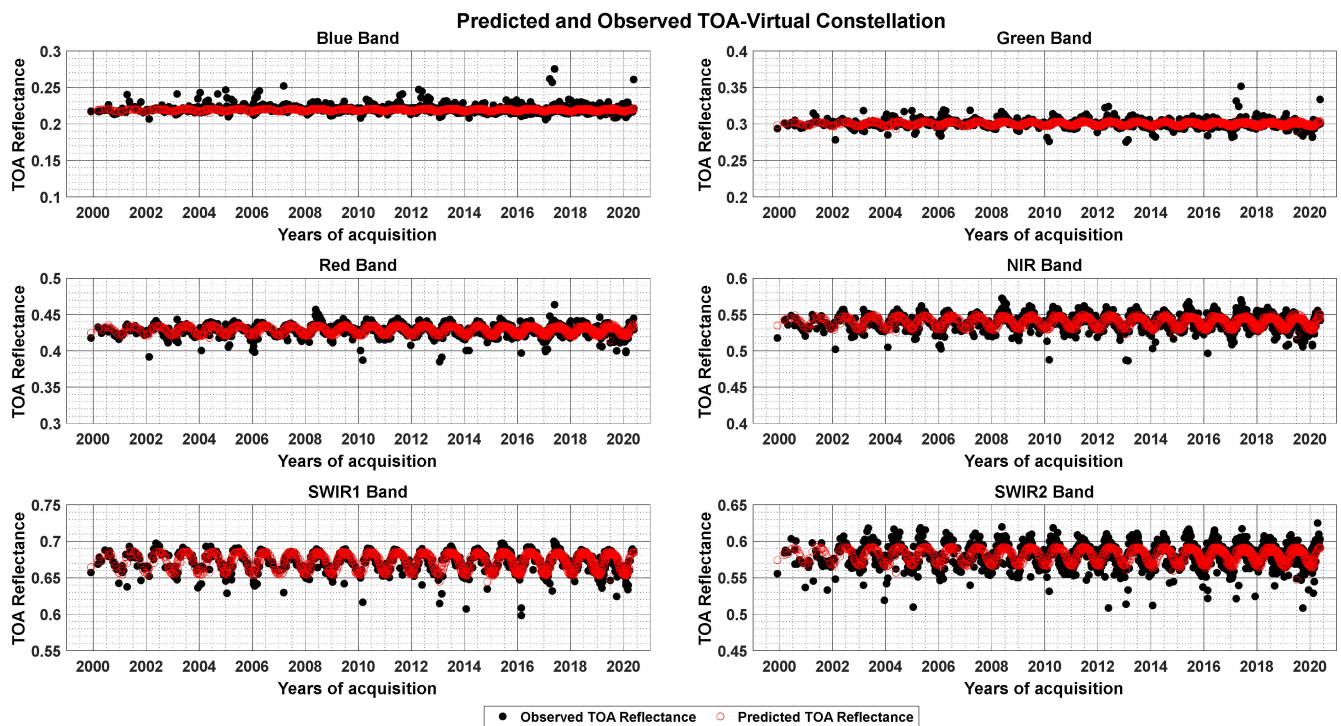

Figure 4. In red, the TOA reflectance predicted by the BRDF model; and in black, the observed TOA reflectance data of virtual constellation over Sudan 1 for blue, green, red, NIR, SWIR 1, and SWIR 2 bands.

Table 5. BRDF model error (residual error) in percentage for all PICS sites in matching spectral bands.

Bands	Sudan 1	Libya 4	Egypt 1	Niger 1	Libya 1	Niger 2
Blue	0.304%	0.259%	0.107%	0.385%	0.398%	0.611%
Green	0.032%	0.059%	−0.017%	0.083%	0.128%	0.281%
Red	−0.113%	−0.043%	−0.056%	−0.045%	0.013%	−0.007%
NIR	−0.116%	−0.083%	−0.055%	−0.055%	−0.014%	−0.056%
SWIR 1	−0.194%	−0.170%	−0.055%	−0.167%	−0.145%	−0.120%
SWIR 2	−0.231%	−0.351%	−0.092%	−0.270%	−0.162%	−0.261%

3.1.3. Inverse Variance Weighted Moving Average Method

Next, the inverse variance weighted moving average method was applied to scaling adjusted BRDF normalized TOA reflectance dataset. Table 6 summarizes the total average estimated uncertainty considering all four sources of uncertainty mentioned in Section 2.4.1. The scaling adjusted BRDF normalized TOA reflectance and the TOA reflectance obtained after applying the weighted moving average trends over Sudan 1 are shown in Figure 5. The weighted averaged TOA reflectance shows no fluctuations in the dataset and highlights the long-term trends. In addition, the residual seasonality seen in the dataset even after BRDF normalization has been minimized.

Table 6. Total average estimated uncertainty for all PICS sites in matching spectral bands.

Bands	Libya 1	Libya 4	Niger 1	Niger 2	Sudan 1	Egypt 1
Blue	6.21%	4.89%	5.43%	5.57%	4.52%	4.59%
Green	4.95%	4.85%	4.24%	4.59%	4.06%	4.68%
Red	4.34%	4.87%	4.00%	4.25%	4.07%	4.63%
NIR	4.47%	5.07%	4.06%	4.33%	4.07%	4.57%
SWIR 1	5.20%	5.54%	4.20%	4.63%	4.25%	4.63%
SWIR 2	5.58%	6.34%	4.95%	5.31%	4.95%	4.82%
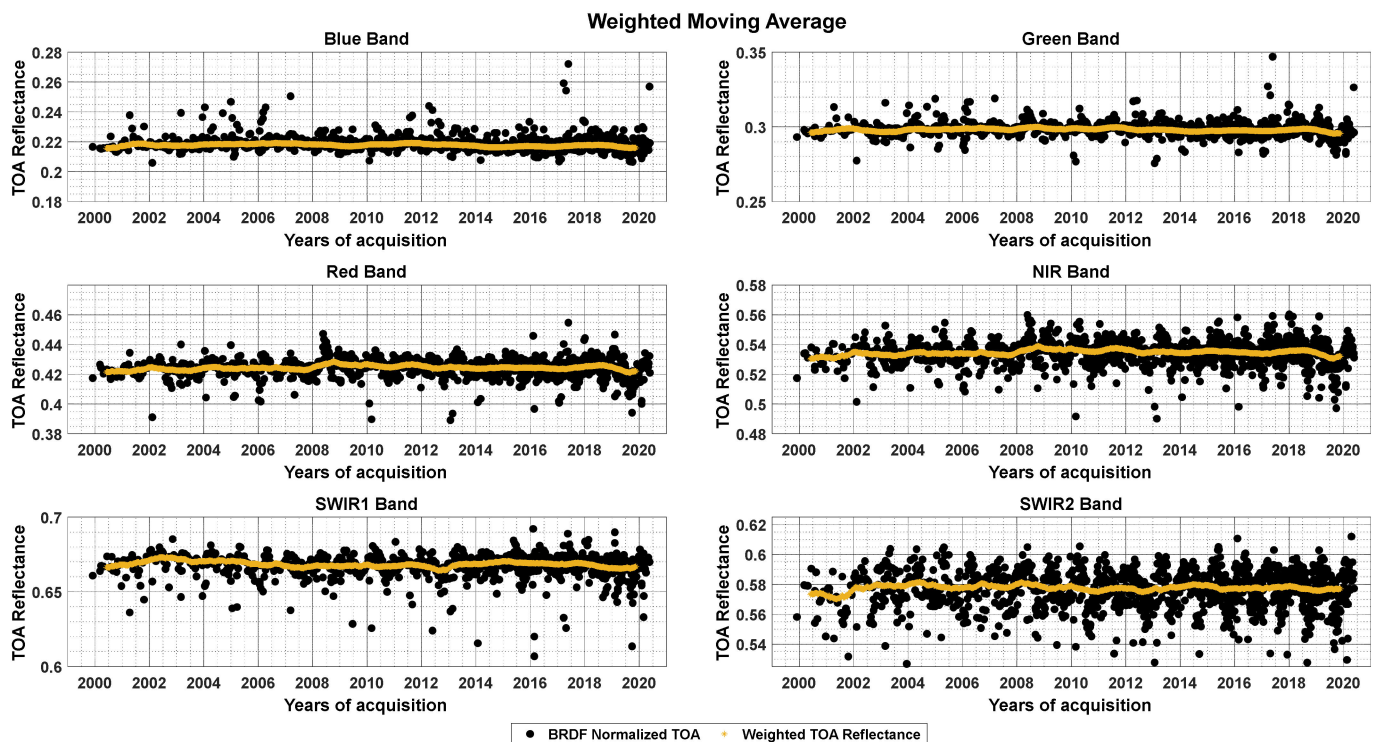
Figure 5. Weighted moving averaged TOA reflectance data of the virtual constellation over Sudan 1 for blue, green, red, NIR, SWIR 1, and SWIR 2 bands.

3.1.4. Sequential Mann–Kendall Test

The results of the sequential Mann–Kendall (SQMK) test statistic for the virtual constellation weighted averaged TOA reflectance dataset of Sudan 1 reveal a statistically significant change point. The yearly plots for the progressive series $u(t_i)$ and retrograde series $u'(t_i)$ for each band are shown in Figure 6. The level of significance was set to 99% (i.e., critical value ± 2.58). For the blue band, the progressive and the retrograde series intersect four times throughout the time series indicating an abrupt change in the years 2007, 2009, 2011, and 2015. For the green band, the change points were detected in the year 2001, 2002, 2009 and 2015. Here, 2015 cannot be recognized as a significant change point as the associated probability value is lower than the accepted level of significance. For the red band, the progressive and the retrograde series cross each other twice in the years 2007 and 2009, indicating a significant turning point. The results for the NIR and red bands are identical. For the SWIR 1 band, the turning point was observed in the years 2002 and 2007. The SQMK result for the SWIR 2 band is also similar. The trend and change behavior of band pairs blue and green, red and NIR, and SWIR 1 and SWIR 2 are expected to be similar since their frequency range along the electromagnetic spectrum to acquire the data are close to each other. Table 7 summarizes change points detected by the SQMK test for all six spectral bands for Sudan 1 site.
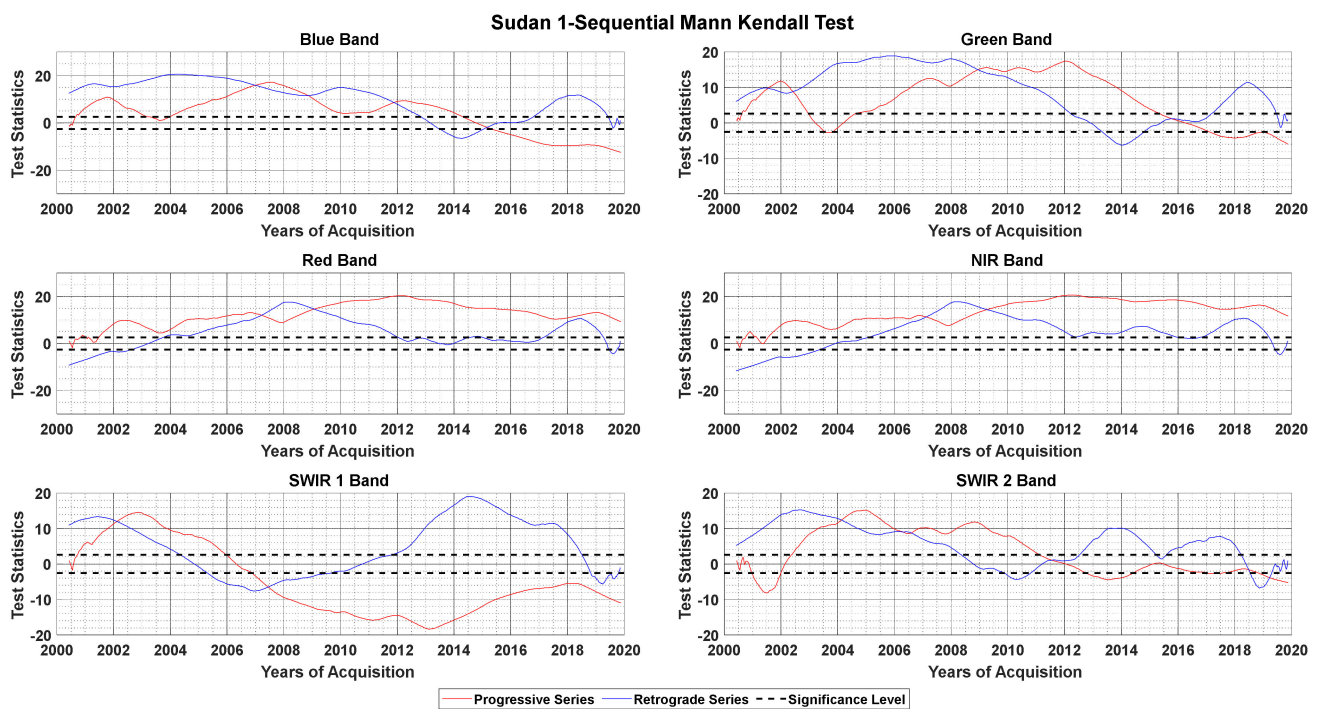
Figure 6. Abrupt change in TOA reflectance data of virtual constellation over Sudan 1 for blue, green, red, NIR, SWIR 1, and SWIR 2 bands.

Table 7. Change point detection by the sequential version of the Mann–Kendall test for Sudan 1.

Bands	Detected Change Points (Year)					Remarks
	1st	2nd	3rd	4th	5th	
Blue	2007 *	2009 *	2011 *	2015 *	-	Significant
Green	2001 *	2002 *	2009 *	2015	-	Significant
Red	2007 *	2009 *	-	-	-	Significant
NIR	2007 *	2009 *	-	-	-	Significant
SWIR 1	2002 *	2007 *	-	-	-	Significant
SWIR 2	2004 *	2007 *	2011	2018	2019 *	Significant

Note: asterisk (*) indicates significant change points at the 99% confidence level.

3.1.5. Mann–Kendall Test

The presence of monotonic trends and their statistical significance was verified using the non-parametric Mann–Kendall (MK test) at a significance level of $\alpha = 0.01$. The MK test was applied to: (a) the entire time series to indicate the long-term trends; and at (b) sub-periods based on change-point results to indicate short-term trends.

The results of the MK test applied to the entire time series to identify a long-term trend is presented in Table 8. The test results indicate that there is a presence of long-term monotonic trends in all the bands. On the other hand, the test was not able to detect a trend in some of the bands (red, NIR, SWIR 1, and SWIR 2) when applied to the unweighted TOA reflectance due to the presence of random variation in the dataset. This suggests that applying the statistical test to a weighted moving averaged dataset gives an advantage of detecting the underlying trend over the unweighted dataset.

Table 8. Mann–Kendall test results at 0.01 significance level for Sudan 1.

Bands	Weighted TOA Reflectance Data		Unweighted TOA Reflectance Data	
	p-Value	Decision	*p*-Value	Decision
Blue	8.2×10^{-36}	Downward Trend	4.3×10^{-16}	Downward Trend
Green	1.4×10^{-9}	Downward Trend	1.6×10^{-8}	Downward Trend
Red	2.5×10^{-20}	Upward Trend	0.258	No Trend
NIR	2.1×10^{-31}	Upward Trend	0.932	No Trend
SWIR 1	9.1×10^{-28}	Downward Trend	0.036	No Trend
SWIR 2	1.5×10^{-7}	Downward Trend	0.594	No Trend

For the descriptive analysis of the site's stability, the MK test was applied to the sub-periods based on the change point results (SQMK test) presented in Section 3.1.4. Figure 7 highlights the significant monotonic trends before and after the turning points throughout the time series. The points in between the trend line represents the inflection point (change points) in the particular band. There is a decreasing trend for the blue band from the year 2007 to 2009, an increasing trend from the year 2009 to 2011, and again, a decreasing trend from the year 2011 to 2015, indicating the years 2007, 2009, 2011, and 2015 as change points. For the green band, the years 2001, 2009, and 2015 were identified as abrupt change points with an increasing trend before the year 2009 and a decreasing trend after that. The red and the NIR bands for this particular site have a change point in the year 2009, with an increasing trend followed by a decreasing trend. SWIR 1 band has a downward trend from 2002 to 2007 and an upward trend from 2007 onwards and the years 2002 and 2007 are the significant change points. The SWIR 2 band has an upward trend at the beginning of the time series and a downward trend from the year 2004 to 2011 and again, upward trend from the year 2011 until 2018.
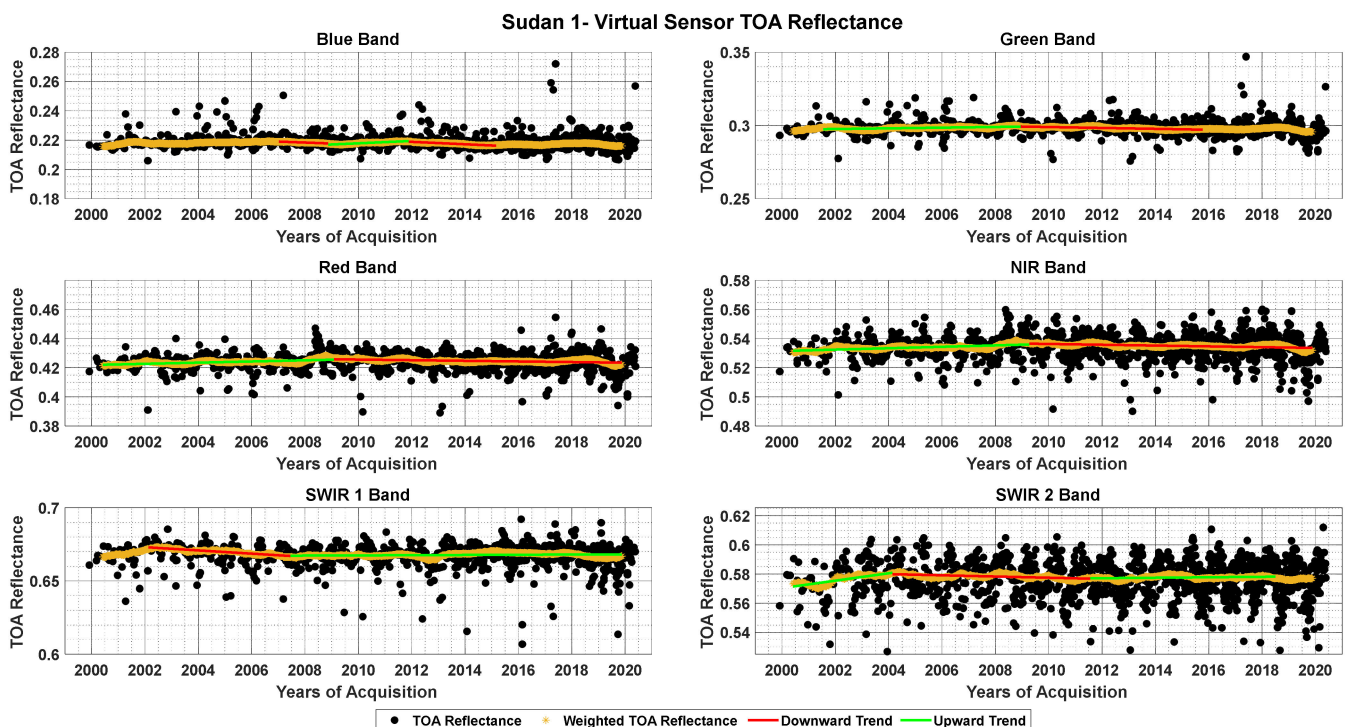
Figure 7. TOA reflectance data of virtual constellation over Sudan 1 for blue, green, red, NIR, SWIR 1, and SWIR 2 bands with fitted trend line before and after turning points.

3.2. Stability Analysis of Libya 4, Egypt 1, Niger 1, Niger 2 and Libya 1

The summary of results obtained for Libya 4, Egypt 1, Niger 1, Niger 2 and Libya 1 are presented in this section.

The MK test results applied to the sub-periods defined by the results of the SQMK test are shown in Figure 8. The figure shows the weighted averaged TOA reflectance dataset divided into sub-periods over Libya 4 for corresponding bands. These sub-periods are fitted with a significant trend line ($\alpha = 0.01$) based on the MK test results both before and after the change points. Figure 8 depicts pronounced upward and downward trends occurring at different time periods throughout the entire time series. The MK test indicates a significant upward trend followed by a downward trend in the blue band, highlighting the actual change points in the years 2009, 2013, and 2018. The green band has a similar trend pattern as the blue band, but the upward trend starts one year earlier with the change points in the years 2008, 2013, and 2017. Although the slope of these trend lines is small ($\sim 10^{-4}$), the inflection points (where the function changes concavity) between these trend lines are statistically significant. The red and the NIR bands also indicate an upward trend and a downward trend with the inflection point in the year 2008 between these trend lines. The inflection points around the year 2012 for SWIR 1 and SWIR 2 can be seen with a downward trend before and upward trend after the inflection point. Table 9 summarizes the change points detected by the SQMK test for all the six spectral bands for the Libya 4 site.
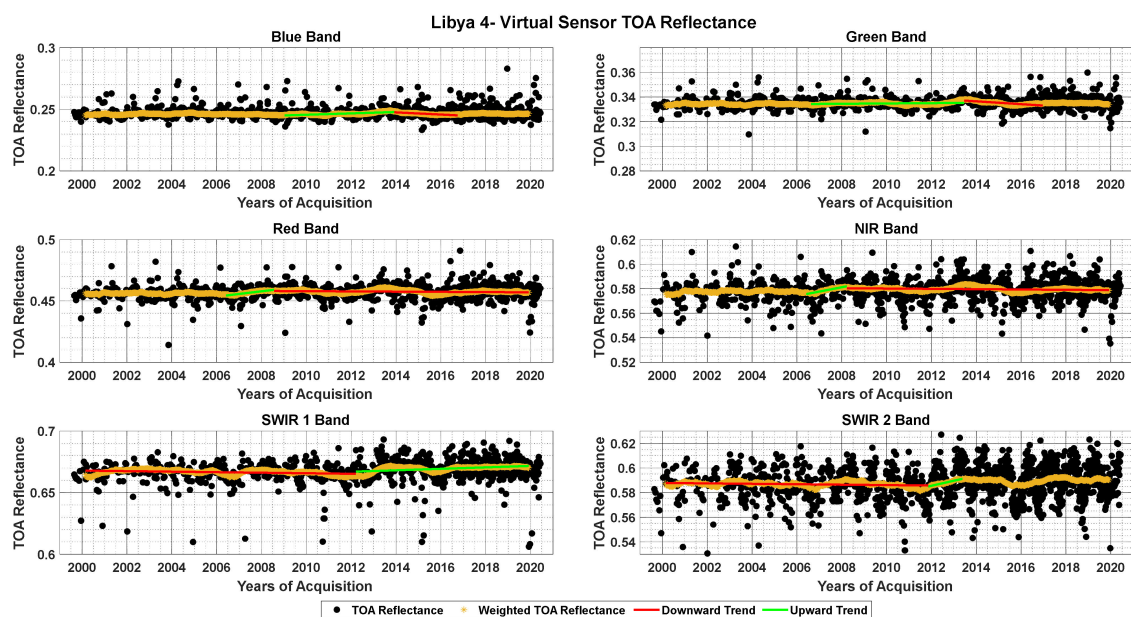

Figure 8. TOA reflectance data of virtual constellation over Libya 4 for blue, green, red, NIR, SWIR 1, and SWIR 2 band with fitted trend lines before and after turning points.

Table 9. Change point detection by sequential version of the Mann–Kendall test for Libya 4.

Bands	Detected Change Points (Year)					Remarks
	1st	**2nd**	**3rd**	**4th**	**5th**	
Blue	2009 *	2013 *	2016	2018 *	-	Significant
Green	2006	2008 *	2011 *	2013 *	2017 *	Significant
Red	2006	2008 *	-	-	-	Significant
NIR	2006	2008 *	-	-	-	Significant
SWIR 1	2012 *	2013 *	-	-	-	Significant
SWIR 2	2011 *	2013 *	2017 *	-	-	Significant

Note: asterisk (*) indicates significant change points at the 99% confidence level.

For Egypt 1, (see Figure 9) there is an increasing trend in the blue and green bands from the beginning of the time series, followed by a decreasing trend towards the end of the time series. However, the change point is detected at different years (2017 and 2019 respectively) for each band. The red and NIR bands have a similar trend, indicating change point in the same year (2006, 2016, and 2019). Although the SWIR bands are also expected to exhibit a similar trend pattern, the test results show contrasting results for this specific site. The decreasing trend is seen at the beginning of the time series until the year 2016 and an increasing trend after that for SWIR 1 band, whereas there is an increasing trend followed by a decreasing trend with the turning point in the year 2016 for SWIR 2 band. The reason behind the contradiction in the trend result for this band pair might be due to the absence of AQUA data (noisy detector) for the SWIR 1 band [41]. Table 10 summarizes the change points detected by the SQMK test for all the six spectral bands for Egypt 1 site.
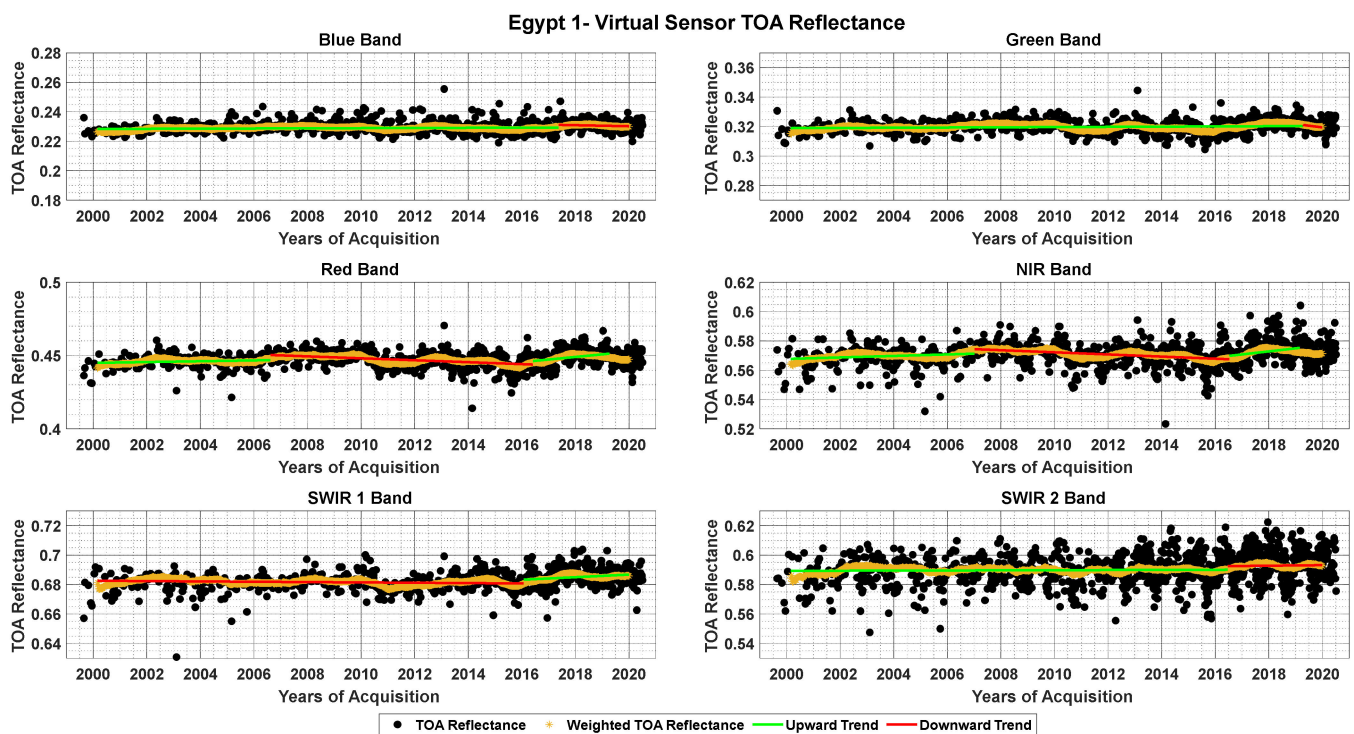

Figure 9. TOA reflectance data of the virtual constellation over Egypt 1 for blue, green, red, NIR, SWIR 1, and SWIR 2 bands with fitted trend lines before and after turning points.

Table 10. Change point detection by the sequential version of the Mann–Kendall test for Egypt 1.

Bands	Detected Change Points (Year)			Remarks
	1st	**2nd**	**3rd**	
Blue	2017 *	2019 *	-	Significant
Green	2017 *	2019 *	-	Significant
Red	2006 *	2017 *	2019	Significant
NIR	2006 *	2017 *	2019	Significant
SWIR 1	2016 *	2018	-	Significant
SWIR 2	2016 *	2018	-	Significant

Note: asterisk (*) indicates significant change points at the 99% confidence level.

For Niger 1, (see Figure 10), the blue band has a single change point in the year 2004 with decreasing trend before and an increasing trend after the change point. The change point in the green band is also detected in the same year (i.e., 2004), but the trend is only seen for a short period (from 2001 to 2003). The red and NIR bands do not have any change point throughout the entire time series. For the SWIR 1 band, there is an increasing trend until the year 2003 and a decreasing trend until the end of the time series. In addition, for SWIR 2 band, there are fluctuations in the trend, indicating a significant change point in the year 2010. Table 11 summarizes the change points detected by the SQMK test for all the six spectral bands for Niger 1 site.
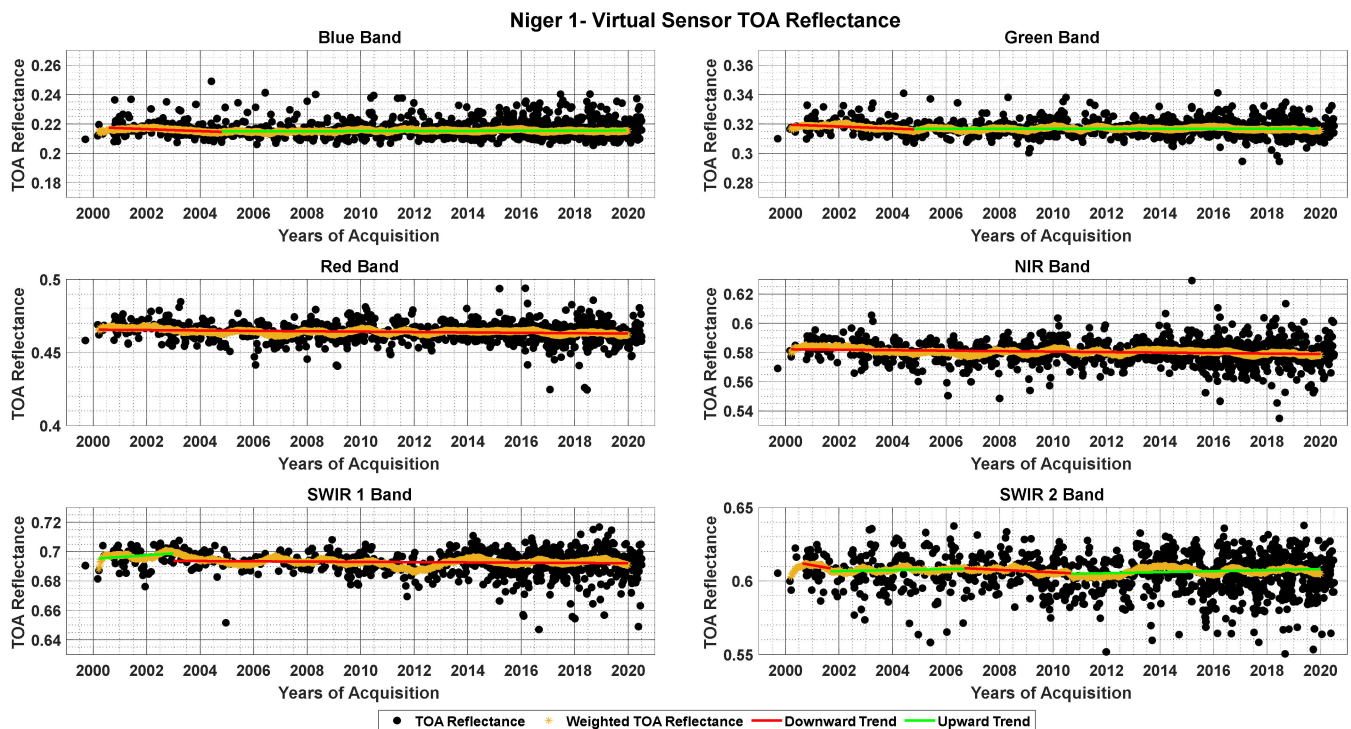

Figure 10. TOA reflectance data of the virtual constellation over Niger 1 for blue, green, red, NIR, SWIR 1, and SWIR 2 bands with fitted trend lines before and after turning points.

Table 11. Change point detection by sequential versions of the Mann–Kendall test for Niger 1.

Bands	Detected Change Points (Year)					Remarks
	1st	**2nd**	**3rd**	**4th**	**5th**	
Blue	2000	2004 *	2016	2018	2019	Significant
Green	2001	2003 *	-	-	-	Significant
Red	-	-	-	-	-	
NIR	-	-	-	-	-	
SWIR 1	2003 *	2004	-	-	-	Significant
SWIR 2	2000	2001 *	2006 *	2010 *	-	Significant

Note: the asterisk (*) indicates significant change points at the 99% confidence level.

For Niger 2, (see Figure 11) it is evident that, like other sites, the band pair blue, green and red-NIR have similar trend patterns and the same year of change point but different trend patterns and same years of the change point for SWIR bands. Table 12 summarizes the change points detected by the SQMK test for all the six spectral bands for the Niger 2 site.
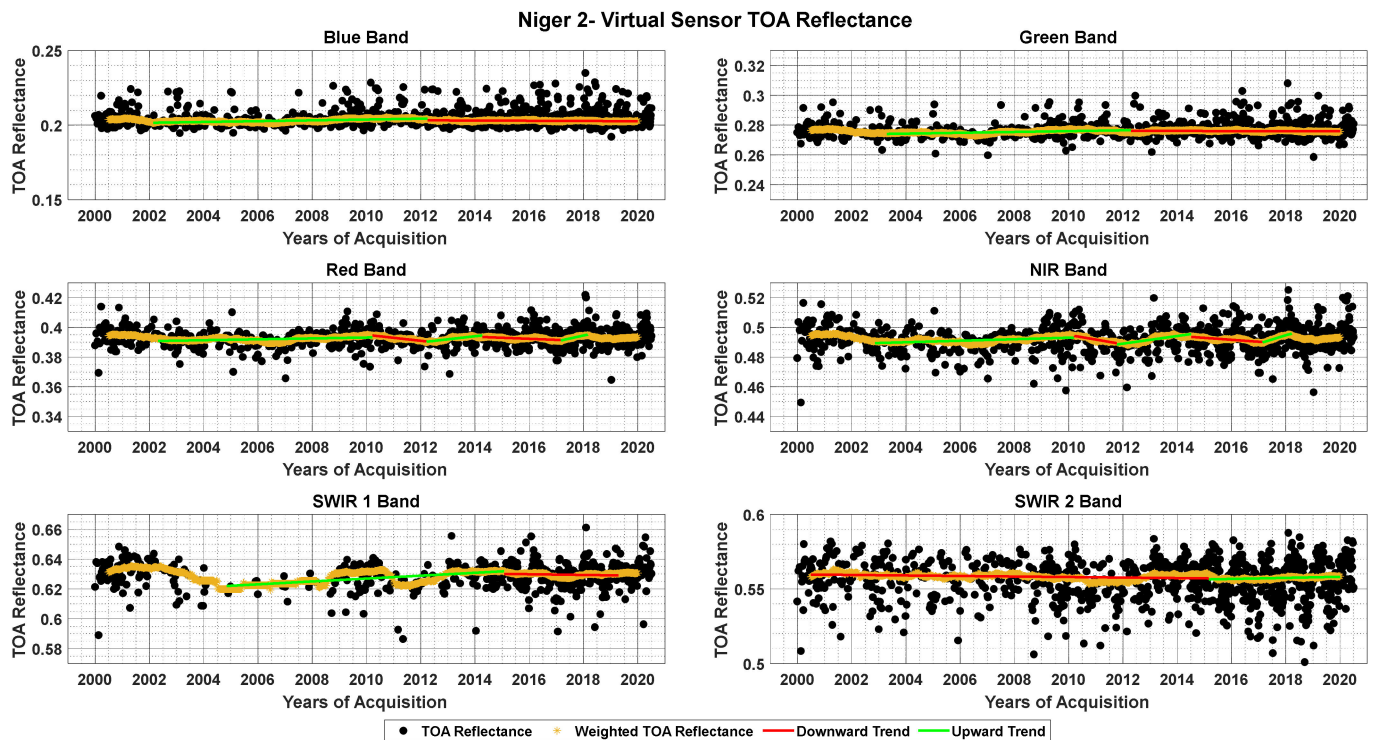

Figure 11. TOA reflectance data of the virtual constellation over Niger 2 for blue, green, red, NIR, SWIR 1, and SWIR 2 bands with fitted trend lines before and after turning points.

Table 12. Change point detection by sequential version of the Mann–Kendall test for Niger 2.

Bands	Detected Change Points (Year)						Remarks
	1st	2nd	3rd	4th	5th	6th	
Blue	2002 *	2012 *	2013	2016	2017	-	Significant
Green	2003 *	2004	2010	2012 *	2015	2017	Significant
Red	2002 *	2010 *	2012 *	2014 *	2017 *	2018 *	Significant
NIR	2002 *	2010 *	2011 *	2014 *	2017 *	2018 *	Significant
SWIR 1	2004 *	2015 *	-	-	-	-	Significant
SWIR 2	2015 *	2016	-	-	-	-	Significant

Note: asterisk (*) indicates significant change points at the 99% confidence level.

For Libya 1, (see Figure 12) the graph for each band indicates the turning point detected by the SQMK test and the trend behavior before and after the turning point. The change points and the trend patterns are again similar for the band pairs blue-green, red-NIR, and SWIR1-SWIR 2. The change points detected by the SQMK test for all the bands are summarized in Table 13.
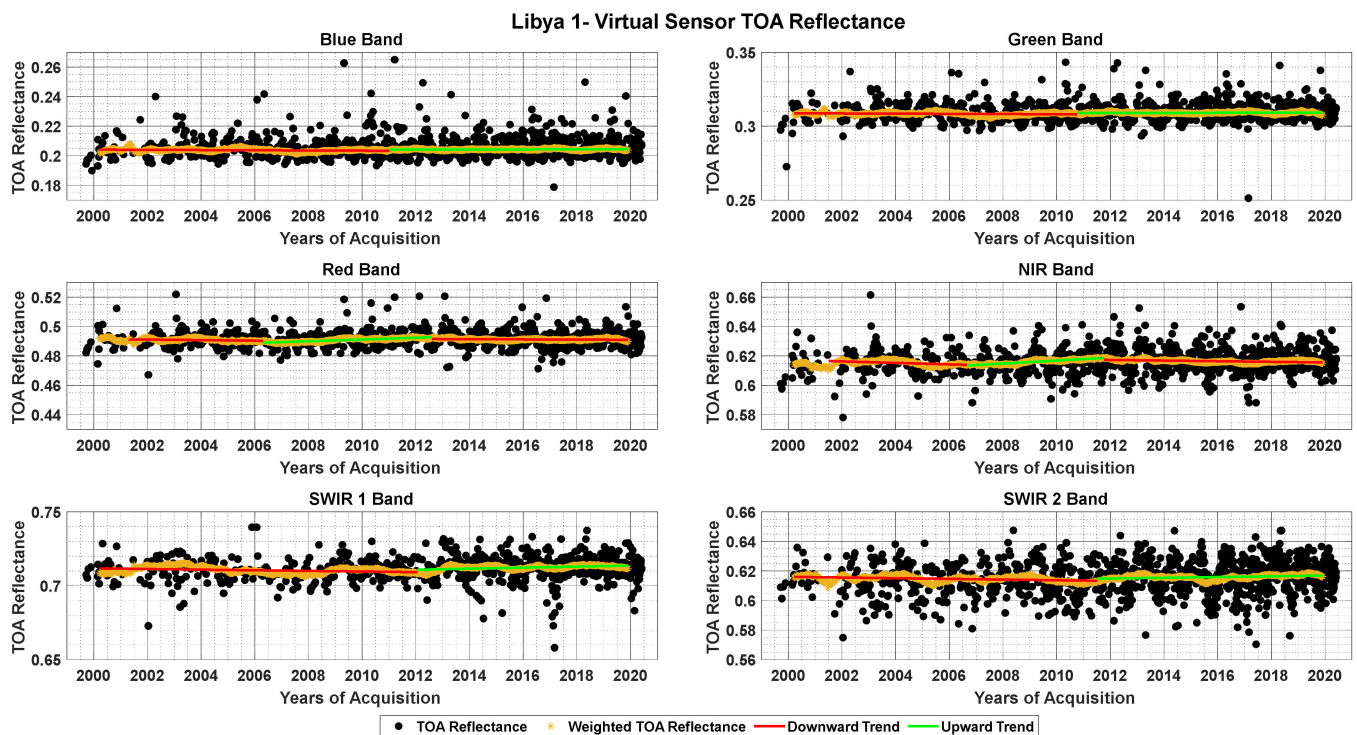
Figure 12. TOA reflectance data of the virtual constellation over Libya 1 for blue, green, red, NIR, SWIR 1, and SWIR 2 bands with fitted trend lines before and after turning points.

Table 13. Change point detection by sequential versions of the Mann–Kendall tests for Libya 1.

Bands	Detected Change Points (Year)					Remarks
	1st	**2nd**	**3rd**	**4th**	**5th**	
Blue	2011 *	-	-	-	-	Significant
Green	2008 *	2011 *	-	-	-	Significant
Red	2001 *	2006 *	2012 *	2018	2019 *	Significant
NIR	2001 *	2006 *	2012 *	-	-	Significant
SWIR 1	2012 *	2014 *	2017 *	-	-	Significant
SWIR 2	2011 *	2012 *	2019 *	-	-	Significant

Note: asterisks (*) indicates significant change points at the 99% confidence level.

The Mann–Kendall test results for the overall long-term trend for Libya 1, Egypt 1, Niger 1, Niger 2 and Libya 1 for all six bands are presented in Table 14. This analysis provides sufficient statistical evidence to state that there is a long-term upward trend in all the bands for Libya 4, Libya 1, and Egypt 1. In contrast, Sudan 1 PICS data indicates a long-term downward trend in all common bands except for Red and NIR bands. Likewise, Niger 1 PICS data indicates a long-term downward trend in all the bands except for the blue band, and Niger 2 indicates temporal stability across all the bands except for green and SWIR 2 bands.

Table 14. Summary of Mann–Kendall test results for overall time series for Libya 1, Egypt 1, Niger 1, Niger 2 and Libya 1.

Bands	*p*-Value	Decision
Libya 4		
Blue	6.32×10^{-12}	Upward Trend
Green	2.85×10^{-20}	Upward Trend
Red	1.64×10^{-35}	Upward Trend
NIR	3.16×10^{-38}	Upward Trend
SWIR 1	9.94×10^{-39}	Upward Trend
SWIR 2	4.08×10^{-42}	Upward Trend
Egypt 1		
Blue	5.28×10^{-44}	Upward Trend
Green	2.73×10^{-14}	Upward Trend
Red	9.86×10^{-11}	Upward Trend
NIR	5.25×10^{-11}	Upward Trend
SWIR 1	5.83×10^{-14}	Upward Trend
SWIR 2	4.42×10^{-31}	Upward Trend
Niger 1		
Blue	0.002	No Trend
Green	5.8×10^{-35}	Downward Trend
Red	3.19×10^{-51}	Downward Trend
NIR	3.4×10^{-61}	Downward Trend
SWIR 1	6.5×10^{-46}	Downward Trend
SWIR 2	1.14×10^{-9}	Downward Trend
Niger 2		
Blue	0.033	No Trend
Green	3.60×10^{-14}	Upward Trend
Red	0.031	No Trend
NIR	0.041	No Trend
SWIR 1	0.532	No Trend
SWIR 2	4.47×10^{-46}	Downward Trend
Libya 1		
Blue	1.82×10^{-25}	Upward Trend
Green	2.19×10^{-23}	Upward Trend
Red	1.69×10^{-7}	Upward Trend
NIR	2.36×10^{-22}	Upward Trend
SWIR 1	1.48×10^{-22}	Upward Trend
SWIR 2	3.89×10^{-11}	Upward Trend

4. Discussion

Taking the advantage of the virtual constellation approach to combine the data from multiple sensors into a single time series and separating sensor changes from site changes, this work focused on quantifying the temporal variability of six PICS. Based on the analysis of the results, it can be inferred that there is a presence of both short-term and long-term monotonic trends in the six common reflective bands of five sensors for all six evaluated PICS. Table 15 summarizes the long-term reflectance change for all the bands over all the six PICS sites. The negative and positive sign in this table indicates a monotonic downward and upward trend, respectively. Even though the trend detected are statistically significant, the magnitude of the slope is noticeably small ranging from 0.008% to 0.215% change in reflectance unit per year. Hence, these trends do not suggest that these sites are changing excessively over time. The findings of the current study and findings of the previous study by Tuli [13] are different in context of presence of long- term trend in some of the sites, which is expected because both the work employs different statistical approaches. For example, the result of the previous study does not indicate any monotonic trend in six

reflective solar bands for Libya 4 and Egypt 1 whereas; the current study shows that there is a presence of long-term upward trend in all the solar reflective bands for Libya 4 and Egypt 1. Likewise, for other sites the current study indicates presence of long-term monotonic trend that previous study was unable to detect. This implies that the current approach is more robust as weighting the TOA reflectance by uncertainty removes the biasness of the statistical test on the data and highlights the underlying long-term trend.

Table 15. Summary of long-term TOA reflectance change per year (in %) for all PICS.

Bands	Sudan 1	Libya 4	Egypt 1	Niger 1	Libya 1	Niger 2
Blue	−0.028%	0.008%	0.042%	-	0.019%	-
Green	−0.011%	0.013%	0.026%	−0.022%	0.015%	0.010%
Red	0.016%	0.019%	0.019%	−0.028%	0.006%	-
NIR	0.019%	0.019%	0.017%	−0.029%	0.015%	-
SWIR 1	−0.018%	0.002%	0.015%	−0.028%	0.019%	-
SWIR 2	−0.009%	0.039%	0.025%	−0.012%	0.011%	−0.215%

The results also provide evidence from a statistical perspective to indicate significant change points in all the sites. The change points have wider implication for those utilizing the calibration sites as they highlight the dynamic nature of the sites. These changes might be atmospheric change, climate change, or changes due to human activities. Therefore, the analysis presented here could be extended further to try to find out the reason behind the change or relate these changes to an actual physical change with the use of independent data and meteorological records. From the results, it is also clear that the amount of temporal change throughout the time series varies at different sub-periods for different spectral bands, but the results on the trend behavior are broadly consistent with the band pairs blue/green, red/NIR, and SWIR1/SWIR 2 in most of the sites.

The findings of this study suggest that, even though some changes are statistically significant, they can still be used for radiometric calibration. However, for the appropriate use of the PICS data, the user should be careful while implying the change observed in the temporal stability of the PICS directly to the sensor response as some amount of changes might be coming from the site itself. The magnitude of the trends (amount of site change) reported in this work can be useful to reduce the calibration uncertainty of the satellite sensors. Another important aspect to take into consideration is the stability requirement of the PICS for each of the satellite sensor missions. For example, the maximum amount of temporal change detected in all the sites was 0.215% per year for Niger 2. This magnitude of change increases with the increase in length of time of change. In 10 years' time, the magnitude of the change becomes 2.15%, which is less than the stated mission requirement (i.e., 5% calibration uncertainty) for Landsat-7 ETM+; therefore, the site change may be neglected. Nevertheless, some sensors demonstrate finer calibration uncertainty. The radiometric accuracy of sensor Traceable Radiometry Underpinning Terrestrial and Helio Studies (TRUTHS) is 0.3% [42]. In this case, the observed temporal change in 10 years would be 3%, which cannot be overlooked, and the site cannot be considered as a viable source of radiometric calibration.

In summary, this analysis seems to improve the understanding of the stability of PICS by locating the change point and revealing the underlying trend that previous study methods were incapable of detecting. It also has outlined the importance of quantifying the temporal stability of PICS, showing that temporal invariance of the site cannot be assumed. This paper may serve as a reference for the calibration engineers who intend to use PICS for the radiometric calibration and performance monitoring of the satellite sensors in the future.

5. Conclusions

Pseudo-invariant calibration sites have been widely used over the last decades for the regular monitoring of the radiometric performance of the Earth-observing satellite sensors. PICS are the regions on the Earth's surface that are assumed to exhibit minimal change over a long period. Historically, any change observed in the temporal stability of the PICS was directly implied to the change that occurred due to the sensor response. However, this assumption has not been assessed extensively in the past. The present study utilizes the virtual constellation approach to form a long-term TOA reflectance time series from five different satellite sensors (Landsat-8 OLI, Landsat-7 ETM+, Sentinel-2A MSI, Terra MODIS, and Aqua MODIS) over six different PICS sites (Libya 4, Sudan 1, Niger 1, Niger 2, Libya 1 and Egypt 1). The underlying trend and the locations where the behavior of PICS data has changed in this time series were identified using the statistical sequential Mann–Kendall test. Finally, the change point determined by the sequential Mann–Kendall test was used to detect short-term trends by applying the Mann–Kendall test to sub-periods of the time series. The results show that the magnitude of weighted averaged TOA reflectance trends for Sudan 1, Libya 4, Egypt 1, Niger 1, and Libya 1 over 20 years are small changing no more than 0.042% per year even though they are statistically significant. For Niger 2, there is a maximum change of 0.215% reflectance unit per year for the SWIR 2 band compared to all other sites that have been evaluated.

The proposed algorithm provides useful results for identifying the stable dataset over a time period to its usefulness for radiometric calibration. It is important to highlight that this approach can be used to detect change points not only for the selected PICS but also for any time-series data that is used by the remote sensing community. Therefore, the analysis presented here could be further extended to determine the temporal stability of other CEOS recommended PICS, extended PICS, and even satellite sensors.

Author Contributions: Conceptualization, N.K., C.T.P. and L.L.; methodology, N.K., C.T.P. and L.L.; software, N.K.; formal analysis, N.K., C.T.P.; investigation, N.K., C.T.P.; resources, C.T.P. and L.L; writing—original draft preparation, N.K.; writing—review and editing, N.K., C.T.P. and L.L.; visualization, N.K.; supervision, C.T.P. and L.L.; project administration, L.L.; funding acquisition, L.L. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not Applicable.

Informed Consent Statement: Not Applicable.

Data Availability Statement: Not Applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Butler, J.J.; Brown, S.W.; Saunders, R.D.; Johnson, B.C.; Biggar, S.F.; Zalewski, E.F.; Markham, B.L.; Gracey, P.N.; Young, J.B.; Barnes, R.A. Radiometric measurement comparison on the integrating sphere source used to calibrate the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Landsat 7 Enhanced Thematic Mapper Plus (ETM+). *J. Res. Natl. Inst. Stand. Technol.* **2003**, *108*, 199. [CrossRef] [PubMed]
2. Yoon, H.W.; Gibson, C.E.; Barnes, P.Y. Realization of the National Institute of Standards and Technology detector-based spectral irradiance scale. *Appl. Optics* **2002**, *41*, 5879–5890. [CrossRef] [PubMed]
3. USGS. Landsat Missions. Available online: https://www.usgs.gov/core-science-systems/nli/landsat/landsat-collection-1-level-1-quality-assessment-band?qt-science_support_page_related_con=0#qt-science_support_page_related_con (accessed on 3 January 2021).
4. Morstad, D.L.; Helder, D.L. Use of pseudo-invariant sites for long-term sensor calibration. In Proceedings of the IGARSS 2008-2008 IEEE International Geoscience and Remote Sensing Symposium, Boston, MA, USA, 22–27 July 2008; pp. I-253–I-256.
5. Govaerts, Y.M.; Clerici, M. Evaluation of radiative transfer simulations over bright desert calibration sites. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 176–187. [CrossRef]

6. Mishra, N.; Haque, M.O.; Leigh, L.; Aaron, D.; Helder, D.; Markham, B. Radiometric cross calibration of Landsat 8 operational land imager (OLI) and Landsat 7 enhanced thematic mapper plus (ETM+). *Remote Sens.* **2014**, *6*, 12619–12638. [CrossRef]

7. Angal, A.; Xiong, X.; Choi, T.; Chander, G.; Mishra, N.; Helder, D.L. Impact of Terra MODIS Collection 6 on long-term trending comparisons with Landsat 7 ETM+ reflective solar bands. *Remote Sens. Lett.* **2013**, *4*, 873–881. [CrossRef]

8. Vuppula, H. Normalization of pseudo-invariant calibration sites for increasing the temporal resolution and long-term trending. Master's Thesis, South Dakota State University, Brookings, SD, USA, 2017.

9. Chander, G.; Angal, A.; Choi, T.; Xiong, X. Radiometric cross-calibration of EO-1 ALI with L7 ETM+ and Terra MODIS sensors using near-simultaneous desert observations. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, *6*, 386–399. [CrossRef]

10. Li, S.; Ganguly, S.; Dungan, J.L.; Wang, W.; Nemani, R.R. Sentinel-2 MSI radiometric characterization and cross-calibration with Landsat-8 OLI. *Adv. Remote Sens.* **2017**, *6*, 147. [CrossRef]

11. Helder, D.; Thome, K.J.; Mishra, N.; Chander, G.; Xiong, X.; Angal, A.; Choi, T. Absolute radiometric calibration of Landsat using a pseudo invariant calibration site. *IEEE Trans. Geosci. Remote Sens.* **2013**, *51*, 1360–1369. [CrossRef]

12. Raut, B.; Kaewmanee, M.; Angal, A.; Xiong, X.; Helder, D. Empirical Absolute Calibration Model for Multiple Pseudo-Invariant Calibration Sites. *Remote Sens.* **2019**, *11*, 1105. [CrossRef]

13. Tuli, F.T.Z.; Pinto, C.T.; Angal, A.; Xiong, X.; Helder, D. New Approach for Temporal Stability Evaluation of Pseudo-Invariant Calibration Sites (PICS). *Remote Sens.* **2019**, *11*, 1502. [CrossRef]

14. Claverie, M.; Ju, J.; Masek, J.G.; Dungan, J.L.; Vermote, E.F.; Roger, J.-C.; Skakun, S.V.; Justice, C. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* **2018**, *219*, 145–161. [CrossRef]

15. CEOS. Virtual Constellations. Available online: http://ceos.org/ourwork/virtual-constellations/ (accessed on 15 October 2020).

16. Wulder, M.A.; Hilker, T.; White, J.C.; Coops, N.C.; Masek, J.G.; Pflugmacher, D.; Crevier, Y. Virtual constellations for global terrestrial monitoring. *Remote Sens. Environ.* **2015**, *170*, 62–76. [CrossRef]

17. USGS. Landsat Missions. Available online: https://www.usgs.gov/core-science-systems/nli/landsat (accessed on 15 October 2020).

18. Barsi, J.A.; Markham, B.L.; Czapla-Myers, J.S.; Helder, D.L.; Hook, S.J.; Schott, J.R.; Haque, M.O. Landsat-7 ETM+ radiometric calibration status. In Proceedings of the Earth Observing Systems XXI, San Diego, CA, USA, 30 August–1 September 2016; p. 9972.

19. ESA. Mission Objectives. Available online: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2/mission-objectives (accessed on 17 October 2020).

20. Gascon, F.; Bouzinac, C.; Thépaut, O.; Jung, M.; Francesconi, B.; Louis, J.; Lonjou, V.; Lafrance, B.; Massera, S.; Gaudel-Vacaresse, A. Copernicus Sentinel-2A calibration and products validation status. *Remote Sens.* **2017**, *9*, 584. [CrossRef]

21. Barsi, J.A.; Alhammoud, B.; Czapla-Myers, J.; Gascon, F.; Haque, M.O.; Kaewmanee, M.; Leigh, L.; Markham, B.L. Sentinel-2A MSI and Landsat-8 OLI radiometric cross comparison over desert sites. *Eur. J. Remote Sens.* **2018**, *51*, 822–837. [CrossRef]

22. Barnes, W.L.; Salomonson, V.V. MODIS: A global imaging spectroradiometer for the Earth Observing System. In Proceedings of the Optical Technologies for Aerospace Sensing: A Critical Review, Boston, MA, USA, 16 November 1992; p. 102690G.

23. Xiong, X.; Sun, J.; Barnes, W.; Salomonson, V.; Esposito, J.; Erives, H.; Guenther, B. Multiyear on-orbit calibration and performance of Terra MODIS reflective solar bands. *IEEE Trans. Geosci. Remote Sens.* **2007**, *45*, 879–889. [CrossRef]

24. Xiong, X.; Sun, J.; Xie, X.; Barnes, W.L.; Salomonson, V.V. On-orbit calibration and performance of Aqua MODIS reflective solar bands. *IEEE Trans. Geosci. Remote Sens.* **2009**, *48*, 535–546. [CrossRef]

25. Mishra, N.; Helder, D.; Angal, A.; Choi, J.; Xiong, X. Absolute calibration of optical satellite sensors using Libya 4 pseudo invariant calibration site. *Remote Sens.* **2014**, *6*, 1327–1346. [CrossRef]

26. Goddard Space Flight Center; Landsat 7 Project Science Office. *Landsat 7: Science Data User's Handbook*; NASA Goddard Space Flight Center; Landsat 7 Project Science Office: Greenbelt, MD, USA, 1999.

27. USGS. *Landsat 8 Data Users Handbook*, 5 ed.; EROS Sioux Falls: Sioux Falls, SD, USA, 2019.

28. Liu, J.-J.; Li, Z.; Qiao, Y.-L.; Liu, Y.-J.; Zhang, Y.-X. A new method for cross-calibration of two satellite sensors. *Int. J. Remote Sens.* **2004**, *25*, 5267–5281. [CrossRef]

29. Schläpfer, D.; Richter, R.; Feingersh, T. Operational BRDF effects correction for wide-field-of-view optical scanners (BREFCOR). *IEEE Trans. Geosci. Remote Sens.* **2014**, *53*, 1855–1864. [CrossRef]

30. Farhad, M.; Kaewmanee, M.; Leigh, L.; Helder, D. Radiometric Cross Calibration and Validation Using 4 Angle BRDF Model between Landsat 8 and Sentinel 2A. *Remote Sens.* **2020**, *12*, 806. [CrossRef]

31. Kaewmanee, M. Pseudo invariant calibration sites: PICS evolution. In Proceedings of the CALCON 2018, Utah State University, Logan, UT, USA, 18–20 June 2018.

32. Tucker, J.; Ortiz, M. Weighting Individual Datum for Nonparametric Analysis. *Proc. Am. Soc. Min. Reclam.* **2007**, 831–836.

33. Bevington, P.R.; Robinson, D.K. *Data Reduction and Error Analysis*; McGraw Hill: New York, NY, USA, 2003.

34. Sneyers, R. *On the Statistical Analysis of Series of Observations*; Secretariat of the World Meteorological Organization: Geneva, Switzerland, 1991.

35. Bisai, D.; Chatterjee, S.; Khan, A.; Barman, N. Application of sequential Mann-Kendall test for detection of approximate significant change point in surface air temperature for Kolkata weather observatory, west Bengal, India. *Int. J. Curr. Res.* **2014**, *6*, 5319–5324.

36. Kundu, A.; Chatterjee, S.; Dutta, D.; Siddiqui, A. Meteorological trend analysis in Western Rajasthan (India) using geographical information system and statistical techniques. *J. Environ. Earth Sci.* **2015**, *5*, 90–99.

37. Chatterjee, S.; Bisai, D.; Khan, A. Detection of Approximate Potential Trend Turning Points in Temperature Time Series (1941-2010) for Asansol Weather Observation Station, West Bengal, India. *Atmos. Clim. Sci.* **2014**, *4*. [CrossRef]

38. Kendall, M.G. *Rank Correlation Methods*; Hafner Publishing Company: Neudruck, NY, USA, 1948.
39. Mann, H.B. Nonparametric tests against trend. *Econom. J. Econom. Soc.* **1945**, 245–259. [CrossRef]
40. Militino, A.F.; Moradi, M.; Ugarte, M.D. On the Performances of Trend and Change-Point Detection Methods for Remote Sensing Data. *Remote Sens.* **2020**, *12*, 1008. [CrossRef]
41. Wang, L.; Qu, J.J.; Xiong, X.; Hao, X.; Xie, Y.; Che, N. A new method for retrieving band 6 of Aqua MODIS. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 267–270. [CrossRef]
42. Gorroño, J.; Banks, A.C.; Fox, N.P.; Underwood, C. Radiometric inter-sensor cross-calibration uncertainty using a traceable high accuracy reference hyperspectral imager. *ISPRS J. Photogramm. Remote Sens.* **2017**, *130*, 393–417. [CrossRef]