This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2024.3376580

A novel inter-calibration method for Fengyun(FY)-3 VIRR using MERSI onboard the same satellite based on pseudo-invariant pixels

Junwei Wang, Xiuqing Hu, *Member, IEEE*, Kun Gao, Yuqing He, Ling Wang, Guorong Li, Na Xu, and Peng Zhang, *Senior Member, IEEE*

***Abstract*—This study presents a novel approach to the radiometric inter-calibration between two sensors onboard the same satellite based on pseudo-invariant pixels (PIPs) using iteratively re-weighted multivariate alteration detection (IR-MAD) method. The IR-MAD algorithm can statistically select pseudo-invariant pixels from the multispectral image pair to assess the radiometric differences between them. Analysis of multiple image pairs from different acquisition times can provide long-term inter-calibration results of the two sensors. The procedure is applied to Fengyun(FY)-3A&3B Visible Infrared Radiometer (VIRR), with the Medium Resolution Spectral Imager (MERSI) onboard the same platform as the reference. Consistency of the spatial distribution of the PIPs selected by IR-MAD with pseudo-invariant calibration sites (PICS) given by other scientists demonstrates the effectiveness of our method. The long-term time series trending of top-of-atmosphere VIRR reflectance over LIBYA1 and LIBYA4 after inter-calibration correction shows that the inter-calibrated VIRR has good agreement with MERSI, with a mean bias of less than 1% and an uncertainty of less than 2% for most channels. The approach requires no prior knowledge of the inter-calibration targets and extends PICS to the pixel-level targets, which results in more diverse samples, broader dynamic ranges and lower uncertainty, yielding consistent and reliable long-term inter-calibration results.**

***Index Terms*—Inter-calibration, Pseudo-invariant pixels, Iteratively reweighted multivariate alteration detection (IR-MAD), Visible Infrared Radiometer (VIRR), Medium Resolution Spectral Imager (MERSI).**

I. Introduction

QUANTITATIVE remote sensing depends on the Earth observing (EO) sensors to provide reliable, accurate, and consistent measurements over time, especially for the long-term trend monitoring of the Earth system. In order to benefit fully from the ever-increasing number of EO satellite systems, inter-calibration between the sensors is critical to bring the measurements from various imaging sensor systems to a common radiometric scale and hence sensor radiometric calibration is of critical importance [1].

The inter-calibration is a technique to use a well-calibrated sensor as a reference to intercalibrate other sensors with near-simultaneous observations of the common targets on the surface of the Earth, Moon or mutual reference to pseudo-invariant features [2]. Consistency biases between different sensors can be introduced from temporal, geometric, and spatial variation in sampling, as well as relative spectral response differences and sensor degradation after launch. Regular inter-calibration is necessary for data continuity and consistency from different imaging sensors, particularly for which without onboard calibrators or where vicarious calibrations are limited. A number of studies have shown that inter-calibration is one of the potential techniques for long-term radiometric trending and quantifying radiometric bias for relative and absolute calibration [3]–[8].

Numerous approaches to inter-calibration between the sensors have been developed and implemented to better quantify the radiometric biases, and new methodologies continue to evolve. They mainly differ in degrees of simultaneity between sensors and the associated ancillary data. The vicarious ground-based calibration method, such as radiance- or reflectance- based methods, rely on simultaneous surface measurements and radiative transfer code computations [9]–[15]. However, these methods typically involve field campaigns, which are cost and labor intensive, hence the number of such calibrations is limited. In an attempt to have more frequent calibration or validation opportunities, certain targets have been used to calibrate or intercalibrate satellite sensors, including pseudo-invariant calibration sites (PICS) [16]–[18], deep convective cloud (DCC) [18], [19], sunglint [20], [21], and the Moon [22], [23]. The simultaneous nadir overpass (SNO) method [3], [24]–[26] was proposed to assess the radiometric consistency between two satellites at the orbital intersections to further reduce uncertainties due to such effects as bi-directional reflectance distribution factor (BRDF). The SNO method was later extended to low latitude (SNO-x) [8], which makes comparisons over deserts and green vegetation possible. These methods generally depend on a series of elaborate thresholds of simultaneity applied to collocate the data of sensors to minimize the consistency biases that may

This work was supported by the National Key Research and Development Program of China [grant numbers 2022YFB3902900, 2022YFB3902901]; the National Natural Science Foundation of China [grant numbers 41871249, U2241275, 61827814], the Beijing Natural Science Foundation [grant number Z190018]; and the China Highresolution Earth Observation System Project [grant number 52-L10D01-0613-20/22].

Corresponding authors: Xiuqing Hu (huxq@cma.gov.cn), Kun Gao (gaokun@bit.edu.cn).

Junwei Wang, Kun Gao, Yuqing He, and Guorong Li are with the School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China and Key Laboratory of Photoelectronic Imaging Technology and System, Ministry of Education, Beijing Institute of Technology, Beijing 100081, China.

Xiuqing Hu, Ling Wang, Na Xu, and Peng Zhang are affiliated with both the Key Laboratory of Radiometric Calibration and Validation for Environmental Satellites, National Satellite Meteorological Center (National Center for Space Weather), China Meteorological Administration, Beijing 100081, China, and Innovation Center for FengYun Meteorological Satellite (FYSIC), Beijing 100081, China.

be attributed to BRDF effects or different contributions of the target spectral signature and atmosphere composition to the spectral response functions (SRFs), and therefore still constrain the opportunities for high quality and frequent sensor calibration and evaluation.

The motivation behind this study is the potential opportunity for an accurate and high frequent long-term inter-calibration for satellite sensors. An iteratively reweighted multivariate alteration detection (IR-MAD) based method is proposed to automatically select pseudo-invariant pixels (PIPs) in the scene for inter-calibration. The PIPs selected by our method are pixel-level, which do not depend on large spatially homogeneous areas, such as PICS. Moreover, the IR-MAD based method can select a greater number of samples with a wider variety of surface types than those selected by SNO or SNO-x method, resulting in a wider dynamic range of reflectance for inter-calibration. The approach requires no prior knowledge of the inter-calibration targets and provide inter-calibration result with high frequency. We describe the method and its implementation on the interclibration of the visible near infrared (VNIR) bands of Visible Infrared Radiometer (VIRR) onboard Fengyun(FY)-3A&B satellites with Medium Resolution Spectral Imager (MERSI) onboard the same platform. Note that this method is also applicable to other situations where the sensors for inter-calibration are onboard different platforms, such as SNO or SNO-x events. Results of inter-calibration between VIRR and MERSI on the same platform demonstrate the efficacy of our method.

II. Sensor Overview and Data

A. Sensor Description

1) FY-3/VIRR: Visible infrared radiometer (VIRR) is a multi-band imager which inherited from FY-1C and FY-1D and continued to be carried onboard FY-3 series sun-synchronous satellites. The involved satellites in this study, *i.e.* FY-3A and FY-3B, are a morning satellite with equator crossing time (ECT) at 10:00 and an afternoon satellite with ECT at 13:30 respectively. VIRR has 10 channels, of which seven visible near infrared (VNIR) channels and three thermal infrared (TIR) channels, with a spatial resolution of 1.1km at nadir for all bands. More details and channel specification is illustrated in [27]. VIRR is not equipped with an onboard calibration system for reflective solar bands (RSBs). The in-orbit test and postlaunch vicarious calibration found that the prelaunch calibration coefficients for the VIRR solar bands are not applicable [28]. The operational calibration depends on the annual site calibration campaign in Dunhuang. However, the operational calibration coefficients are not updated annually, thus the accurate and frequent calibration for the VIRR RSBs is necessary.

2) FY-3/MERSI: Medium Resolution Spectral Imager (MERSI) is the keystone payload, which is completely new generation imager of FY-3 series satellites. MERSI has 20 spectral bands, of which 19 are RSBs and one is TIR band, covering the visible, near-infrared, and thermal infrared spectra. MERSI scan the Earth through a 45° scan mirror in concordance with one K-mirror (derotation), resulting in a swath of 2900km cross-track by 10km along track (at nadir) for each scan [29]. The spatial resolution at nadir of bands 1-5 is 250m, whereas 1000m for the remaining 15 bands. See [27] for more details and channel specification of MERSI. The MERSI is equipped with a visible onboard calibrator (VOC), which is a 6-cm-diameter integrating minisphere designed to monitor the system radiometric response changes which arise either from the MERSI degradation or a change in the output of VOC. However, due to significant degradation of VOC itself, it is not used for updating calibration coefficients on orbit [29]. In practice, the operational calibration coefficients of MERSI are updated based on the vicarious calibration using global multisites method and field measurement campaigns conducted in China radiometric calibration sites, i.e. Dunhuang. Many vicarious calibration methods have been conducted to MERSI and the overall uncertainty in the MERSI top-of-atmosphere radiance or reflectance is less than 5% [29]. In this study, the long-term degradation of MERSI is monitored and corrected using the method in [30], and the results align consistently with other vicarious calibration methods.

B. Study Area and Data

The region of interest (ROI) for this study is located in the North Africa, as shown in Figure 1. The reasons for selecting this region as ROI are as follows: 1) this region is mainly made up of desert with almost no vegetation, because a high reflectance can reduce uncertainties from the atmospheric path radiance due to higher signal-to-noise ratio; 2) this region is arid to minimize the influence of atmospheric water vapor and has minimal cloud cover and precipitation; 3) this region is relatively spatial uniform to minimize the effects of misregistration in inter-calibration; 4) the surface of this region is relatively spectrally uniform, which is particularly important for the matching spectral bands that have different spectral response profiles in inter-calibration; 5) Several reference pseudo-invariant calibration sites (PICS) in this region, such as LIBYA1 and LIBYA4, can be used to evaluate the inter-calibration accuracy.

In this study, the L1B level data of MERSI and VIRR are used for inter-calibration. Experiments were carried out on FY-3A and FY-3B respectively to demonstrate the efficacy and generality of the proposed method. The range of the data covers almost the entire life cycle of the satellite, as illustrated in Table I.

TABLE I
THE TIME RANGE OF FY-3A AND FY-3B DATA USED IN THIS STUDY.

	Start date	End date	Time span
FY-3A	Nov. 12, 2008	Dec. 31, 2014	>6 years
FY-3B	Jan. 21, 2011	Nov. 14, 2018	≈8 years

III. Methodology

A. Inter-calibration Formulation

Although the aperture spectral radiance is actually measured by the sensor, three advantages of converting the at-sensor spectral radiance to top-of-atmosphere (TOA) reflectance were

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2024.3376580

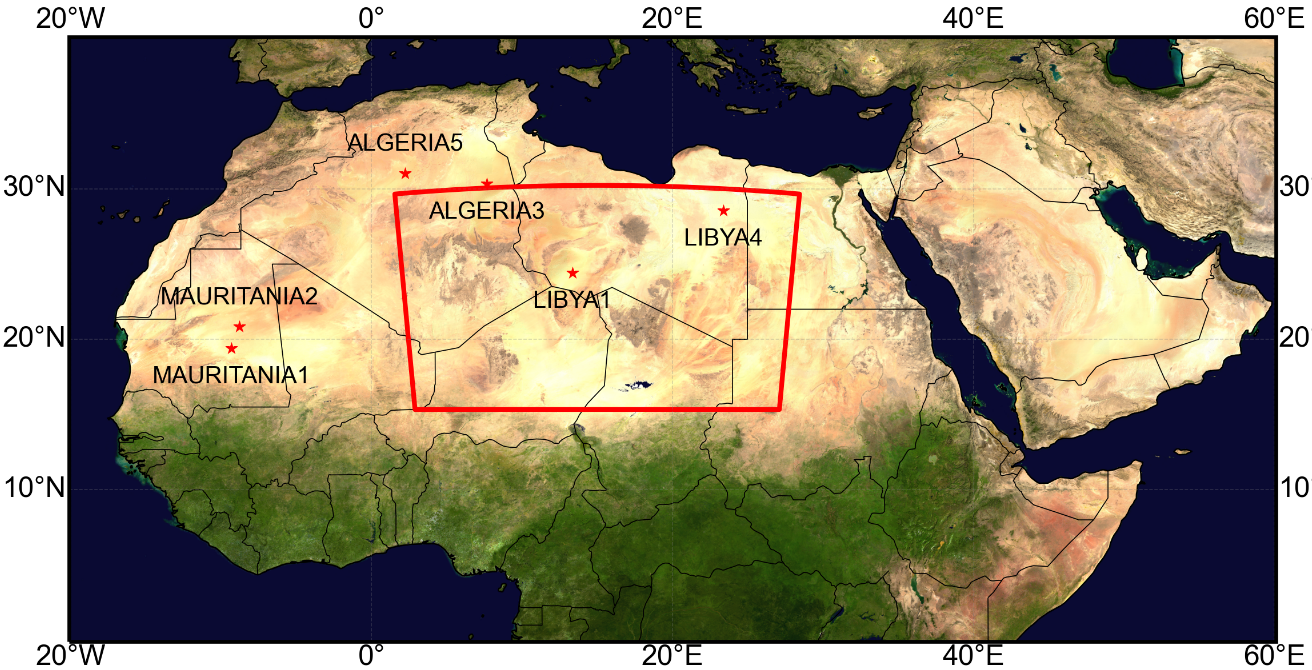

Fig. 1. The ROI in this study. The six reference PICS defined by CEOS are marked as star, of which LIBYA1 and LIBYA4 are located within the ROI.

reported in [31]. Using TOA reflectances instead of radiance can reduce the image-to-image variability and is a fundamental step to bring image data from multiple sensors and platforms to a common scale. For a Lambertian surface in spectral band i, the TOA reflectance can be computed as follows:

$$\rho_i = \frac{\pi d^2 L_i}{E_{0i}\cos\theta} = \frac{\pi}{E_{0i}}\frac{d^2}{\cos\theta}\left(DN_i - C_{0i}\right) S_i \tag{1}$$

where L_i is the spectral radiance at the sensor's aperture [in W/(m^2 sr μm)], E_{0i} is the mean exo-atmospheric solar irradiance [in W/(m^2 μm)] that can be obtained by convolving the solar spectra [32] with the SRF of a given instrument. d is the Earth–Sun distance in astronomical units (AU), θ is the solar zenith angle. DN_i represents raw digital number (in counts) recorded at the satellite, C_{0i} the zero-radiance response (in counts), and S_i the sensor sensitivity coefficient in units of percent reflectance per unit count.

In this study, MERSI is used as the reference sensor to intercalibrate VIRR onboard the same platform. Suppose MERSI is well-calibrated, the consistency bias between the two sensors come from their SRF differences, calibration differences and the VIRR degradation after launch. In this context, the Equation 1 can be expressed separately for image data from the MERSI ("M") and for image data from the VIRR ("V") as follows:

$$\rho_{Mi} = \frac{\pi d^2 L_{Mi}}{E_{0Mi}\cos\theta} = \frac{\pi}{E_{0Mi}}\frac{d^2}{\cos\theta}\left(DN_{Mi} - C_{0Mi}\right) S_{Mi} \tag{2}$$

$$\rho_{Vi} = \frac{\pi d^2 L_{Vi}}{E_{0Vi}\cos\theta} = \frac{\pi}{E_{0Vi}}\frac{d^2}{\cos\theta}\left(DN_{Vi} - C_{0Vi}\right) S_{Vi} \tag{3}$$

where i is the spectrally matching band of MERSI and VIRR. In practice, a fixed sensitivity coefficient from the early operational stage is adopted to calculate the nominal TOA reflectance. Accordingly, Equations 2 and 3 can be rewritten as:

$$\begin{aligned}\rho^*_{Mi} &= \frac{\pi}{E_{M0}}\frac{d^2}{\cos\theta}\left(DN_{Mi} - C_{0Mi}\right) S_{Mi0} \\ &= \frac{\pi}{E_{M0}}\frac{d^2}{\cos\theta}\left(DN_{Mi}\cdot\alpha_{Mi} + \beta_{Mi}\right)\end{aligned} \tag{4}$$

$$\begin{aligned}\rho^*_{Vi} &= \frac{\pi}{E_{V0}}\frac{d^2}{\cos\theta}\left(DN_{Vi} - C_{0Vi}\right) S_{Vi0} \\ &= \frac{\pi}{E_{V0}}\frac{d^2}{\cos\theta}\left(DN_{Vi}\cdot\alpha_{Vi} + \beta_{Vi}\right)\end{aligned} \tag{5}$$

Where ρ^*_{Mi} and ρ^*_{Vi} are nominal TOA reflectance calculated by the fixed sensitivity coefficients S_{Mi0} and S_{Vi0}. Since VIRR has degradation over time, it is useful to further separate the sensitivity coefficient into a fixed initial component and a time varying component as:

$$S_{Vi} = S_{Vi0}\cdot S_{Vi}\left(t\right) \tag{6}$$

where t represents days since the first day (S_{Vi0}). $S_{Vi}\left(t\right)$ is the inverse of the relative degradation. Note that the MERSI is supposed to be well-calibrated, therefore:

$$S_{Mi} = S_{Mi0}\cdot S_{Mi}\left(t\right), S_{Mi}\left(t\right) = 1 \rightarrow S_{Mi} = S_{Mi0} \tag{7}$$

The TOA reflectance of MERSI can be compensated with the spectral band adjustment factor (SBAF) f_{SBAF} and the relative calibration factor f_0, where f_{SBAF} accounts for SRF

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2024.3376580

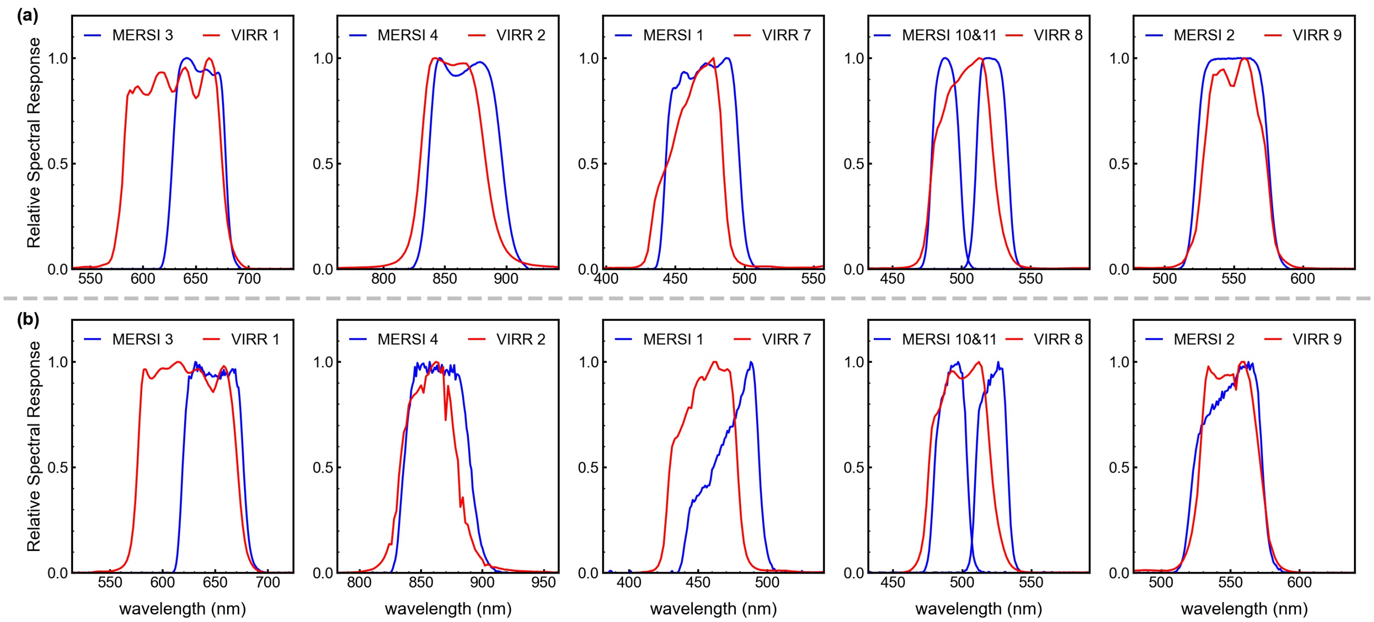

Fig. 2. The relative spectral response functions (SRFs) of matching channels between VIRR and MERSI for (a) FY-3A and (b) FY-3B.

differences and f_0 explains calibration differences for the fixed sensitivity coefficients adopted in Equation 4 and 5. Then,

$$\rho_{Vi} = f_0 \cdot f_{SBAF} \cdot \rho_{Mi} \tag{8}$$

The combination of Equation 2 to 8 yields

$$\rho^*_{Vi} = \frac{f_0 \cdot f_{SBAF}}{S_{Vi}(t)} \rho^*_{Mi} \tag{9}$$

In practice, the α_i and β_i can be read directly from the first day L1B file for MERSI and VIRR. The f_0 we want to obtain is contained in the slope of linear equation that characterizes ρ^*_{Vi} as a function of ρ^*_{Mi}. Once we get the degradation rate function $S_{Vi}(t)$ and relative calibration factor f_0, the updated value of VIRR sensitivity coefficients is then given by

$$S_{Vi} = \frac{S_{Vi0} \cdot S_{Vi}(t)}{f_0} = \frac{\alpha_{Vi} \cdot S_{Vi}(t)}{f_0} \tag{10}$$

With this updated value of S_{Vi}, users can obtain TOA reflectance of VIRR from Equation 3.

B. Spectral Band Matching

Different sensors have varying channels and spectral coverage ranges. Even if two sensors have similar spectral ranges in a given channel, differences in their relative spectral responses (RSRs) may still exist. Figure 2 displays the spectral response functions of matching channels between VIRR and MERSI. The differences in RSRs can lead to systematic biases in measurements of the same radiation source. Therefore, in inter-calibration, the differences caused by the RSRs differences between the two sensors need to be addressed. This can be compensated for by using the SBAF. The definition and calculation of SBAF are documented in [33].

In this study, the Scanning Imaging Absorption Spectrometer for Atmospheric Chartography (SCIAMACHY) hyperspectral data were used to calculate the SBAF for each matching channel between VIRR and MERSI. The spectral response function was convolved with the hyperspectral data, and the SBAF values were obtained by performing linear regression on all samples. A total of 44,511 spectral samples were used, representing one full year of SCIAMACHY data within the ROI and encompassing the spectra of all surface types within the ROI. For VIRR channel 8, a binary linear regression was performed since it matches with two MERSI channels. The SBAF values, center wavelength (CW), correlation coefficient (r), and residual standard deviations for each matching channel between VIRR and MERSI are presented in Table II. The regression plots for SBAF can be found in the supplementary materials. Note that there may also include cloudy spectral samples. Examination of the regression plot reveals that there are no outliers, which can be verified by the residual standard deviations in Table 1. The slightly higher residual standard deviation of VIRR band 1 is primarily attributed to the substantial difference in spectral bandwidth between VIRR band 1 and MERSI band 3 as shown in Fig.2. As stated in [33], narrowband RSRs are more sensitive to changes in the spectrum, leading to a more pronounced effect on SBAF compared to the wideband sensor RSR. For the matched bands of VIRR and MERSI, they are all broad bands with bandwidth greater than 50nm and have weakly gaseous absorption, resulting in the SBAF being less sensitive to the spectral variability of the samples. Hence the impact of cloudy samples on the SBAF regression is negligible.

C. IR-MAD Method for Inter-calibration

The multivariate alteration detection (MAD) technique was first proposed for change detection by [34]. this technique has been used in the automatic normalization of remote sensing images [35], and also for the radiometric calibration of AVHRR reflective bands [36]. An iteratively re-weighted

TABLE II
THE MATCHING CHANNELS BETWEEN VIRR AND MERSI ON FY-3A AND FY-3B, TOGETHER WITH THE SBAF AND REGRESSION RESULTS CALCULATED BY USING SCIAMACHY DATA.

satellite	VIRR		MERSI		SBAF	r	residual
	band	CW	band	CW			std(%)
FY3A	1	0.630	3	0.650	0.9557	0.9973	1.31
	2	0.865	4	0.865	0.9867	0.9998	0.32
	7	0.455	1	0.470	0.9939	0.9999	0.51
	8	0.505	10	0.490	0.6371	0.9999	0.36
			11	0.520	0.3681		
	9	0.555	2	0.550	1.015	0.9990	1.15
FY3B	1	0.630	3	0.650	0.9650	0.9969	1.44
	2	0.865	4	0.865	0.9921	0.9998	0.33
	7	0.455	1	0.470	0.9863	0.9995	0.94
	8	0.505	10	0.490	0.6846	1.0000	0.15
			11	0.520	0.3114		
	9	0.555	2	0.550	0.9993	1.0000	0.02

modification of the MAD transformation (IR-MAD) has been introduced [37] and was extended to radiometric normalization with substantial improvement [38]. This method was later applied to the selection of pseudo-invariant calibration sites in Northwest China [39]. In this study, the IR-MAD technique is used to statistically selected pixel-level targets for inter-calibration. Next, the basic principles of IR-MAD and how it is applied to inter-calibration are explained.

The MAD method can be used to automatically select invariant pixels for multispectral satellite imagery. In the context of radiometric normalization, the invariant pixels for bitemporal image refer to those that are temporally invariant across all spectral bands data during the acquisition time interval. The MAD transformation is linear scale invariant under affine transformations of either or both of the original multispectral images, which is explicitly demonstrated in [35]. For inter-calibration, as shown in Equation 9, there is a linear relationship between the nominal TOA reflectance of the two sensors in matching bands. Given the invariance property of the MAD transformation, it is reasonable to use MAD algorithm to select PIPs that are suitable for inter-calibration. Instead of temporally invariant pixels, the PIPs here refer to the pixels that have linear scale invariance under two sensors with different spectral response characteristics in matching bands. Specifically, the PIPs represent the pixels that conform to the linear relationship in all matching bands under the differences caused by a combination of the relative spectral response characteristics of the two sensors, spectral signature of the target, and the atmospheric composition during overpass. It is intuitively conceivable that when the SRF difference of the two sensors is significant, only targets with relatively smooth spectra profiles will be selected as PIPs. Similarly, when one sensor contains atmospheric absorption feature within the SRF and the other does not, only targets that are not susceptible to atmospheric influence are selected as PIPs.

Consider two sensors with N matching bands for inter-calibration. A image pair of two observations of the common targets from the two sensors can be represented by a random vector $\mathbf{F} = (F_1, ..., F_N)^{\mathrm{T}}$ and $\mathbf{G} = (G_1, ..., G_N)^{\mathrm{T}}$, respectively. [34] proposed that the MAD variates can be determined by a linear transformation of $\mathbf{F}$ and $\mathbf{G}$ with coefficients vector $\mathbf{a}$ and $\mathbf{b}$, and the maximum variance of the MAD variates is achieved.

$$U = \mathbf{a}^{\mathrm{T}}\mathbf{F} = a_1F_1 + a_2F_2 + ... + a_NF_N \tag{11}$$

$$V = \mathbf{b}^{\mathrm{T}}\mathbf{G} = b_1G_1 + b_2G_2 + ... + b_NG_N \tag{12}$$

$$MAD_i = U_i - V_i = \mathbf{a}_i^{\mathrm{T}}\mathbf{F} - \mathbf{b}_i^{\mathrm{T}}\mathbf{G}, i = 1,...,N \tag{13}$$

where the coefficients vector $\mathbf{a}$ and $\mathbf{b}$ can be resolved by applying standard Canonical Correlation Analysis (CCA) [40]. There are some underlying properties for MAD variates: 1) From the Central Limit Theorem, the MAD variates, determined by several additions and subtractions, would ideally fit a normal distribution; 2) Since MAD variates are orthogonal (uncorrelated), all the MAD variates should follow a multivariate normal distribution with diagonal covariance matrix [38].

Let the random variable Z represents the sum of squares of standardized MAD variates:

$$Z = \sum_{i=1}^{N} \left(\frac{MAD_i}{\sigma_i}\right)^2 \tag{14}$$

where σ_i is the variance of MAD_i. Then, Z should follow a chi-square distribution with N degrees of freedom ($\chi_N^2(z)$). An iteration scheme is adopted by setting the probability of no change of observations as weight for the next MAD transformation. The probability of no change of observations z can be determined by the chi-square distribution as follows:

$$P_{no\_change}(z) = 1 - \chi_N^2(z) \tag{15}$$

The general idea behind this formulation is that a small z implies a high probability of no change, resulting in a large weight in the next iteration. This can be considered as more emphasis placed on establishing a better background for detecting change against a background of no change, therefore resulting in improved sensitivity of the MAD transformation [38].

The iteration of MAD transformation will continue until one of the following conditions is met: 1) Maximum number of iterations reached, usually set to 30; 2) The largest absolute change in the canonical correlations, *i.e.* correlations of U and V, becomes smaller than some preset small value (e.g., 10^{-6}). Once the iteration ceases, a decision threshold k can be made to choose the final PIPs for inter-calibration. Typically, the k is set to be the value of z when $P_{no\_change}(z) = 0.9$, that is, pixels which satisfy $P_{no\_change}(z) > 90\%$ are designated as PIPs.

The PIPs are selected statistically from the image pair without a priori knowledge of the target pixel. They should correspond to truly invariant targets for which the overall differences between the image pair can be attributed to linear effects as expressed in Equation 9. The location of the PIPs are likely to change with each image pair, which is reasonable because whether a target is designated as PIP is affected by the atmospheric condition as well as the BRDF effect besides its own spectral signature.

With the selected PIPs, an orthogonal, as opposed to ordinary, linear regression can be performed on the PIPs as demonstrated by [35]. The regression slope $m_i(t)$ provides

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2024.3376580

a measurement of $\frac{f_0 \cdot f_{SBAF}}{S_{Vi}(t)}$ as shown in Equation 9. By performing IR-MAD procedure on image pairs from different acquisition times, a large database of $m_i(t)$ can be created. Analysis of these data can provide long-term inter-calibration results of the two sensors, which can be seen in section V.

In summary, the steps involved in IR-MAD for inter-calibration are as follows:

1) Reproject the near-simultaneous overpass at the ROI from the two sensors onto a common geographic grid to get the image pair.
2) Start with the original MAD transformation for the image pair, i.e., set weights = 1 for all pixels.
3) Iterate the MAD procedure until termination conditions are met:
    a) Set the probability of no change from the last MAD procedure as weight for all pixels.
    b) Perform MAD procedure on the re-weighted image pair.
    c) Calculate weights for the next iteration.
4) Select PIPs from the last MAD procedure by the preset decision threshold k.
5) Perform an orthogonal regression on PIPs to get regression slope m_i, which is a measurement of the inter-calibration result.

IV. Implementation and Analysis

A. Examples Applied on FY-3 VIRR

The long-term time series datasets of MERSI and VIRR onboard FY-3A and FY-3B satellites are created by reprojecting the L1B data onto a common geographic grid in 1km spatial resolution via the nearest neighbor method. This is a prerequisite for making pair-wise comparison for IR-MAD procedure. A rough threshold-based cloud detection algorithm is applied to remove cloud pixels. Nevertheless it must be acknowledged that reduction of variations in the scene can improve the sensitivity of MAD procedure, the IR-MAD method statistically and iteratively selects the truly PIPs hence delicate cloud detection is not necessary.

Our method assumes that radiometric difference between the PIPs from an image pair is solely due to the linear effects as shown in Equation 9, thus other possible causes of change, such effects as BRDF, misregistration and etc., need to be eliminated or at least minimized. Since the two sensors are onboard the same platform, the difference due to BRDF effect is negligible. In order to reduce the uncertainty introduced from misregistration error, only pixels with view zenith angle (VZA) $< 30°$ are used for statistic analysis of IR-MAD and pixels with VZA $< 15°$ are selected as PIPs for orthogonal regression. This is mainly because large geolocation error exists and spatial size of pixels increases at the edge of swath. The impact of different VZAs and scattering regimes was examined in [36].

Once the image pair data are masked as described above, the IR-MAD procedure is applied subsequently to determine the set of PIPs, of which the number can up to several thousand. The number of PIPs of each image pair varies with the present atmospheric and surface conditions. In order to exclude image pairs strongly affected by atmosphere, image pairs with PIPs number less than 1000 or the regression correlation less than 0.95 will be removed. This means that the differences between these image pairs no longer conform to the linear effects due to atmospheric or surface influences and therefore is not suitable for inter-calibration analysis.

Figure 3 shows an implementation of IR-MAD method applied to FY-3B for inter-calibration on January 23, 2011. Figures 3(a) and (b) show the true color images of MERSI and VIRR respectively. The image of MERSI looks a little redder than that of VIRR, which is mainly caused by the difference in SRFs. Figures 3(c) present the mask used before IR-MAD procedure, and the spatial distribution of selected PIPs is shown in Figures 3(d). Interestingly, it can be seen that the CEOS defined PICS, LIBYA1, is automatically selected as PIPs. This suggests that LIBYA1 is indeed a very suitable PICS for inter-calibration, and also proves the efficacy of our method to accurately select PIPs. A potential advantage of using IR-MAD to select targets is that besides the traditional sites used for calibration, which generally have few surface types and limited reflectance dynamic range, the IR-MAD can automatically select thousands of PIPs with various spectral signatures over a wider radiance dynamic range including the lower reflectance targets (see the spatial distribution of PIPs and their corresponding surface types in Figures 3).

PIPs selected by IR-MAD method can then be used in linear regression to obtain the slope m_i for all bands, which is a measurement of the calibration difference between MERSI and VIRR. Figure 4 shows the regression result of the PIPs in Figure 3. Note that the MERSI nominal TOA reflectance (ρ_M^*) has been compensated by f_{SBAF} for comparison. The regression intercept is due to the unaccounted for change in the zero-radiance response of VIRR after launch, which makes the pre-launch coefficients unable to accurately measure its zero-radiance response on orbit. Though the PIPs are spatially dispersed and consist of a variety of surface targets, the correlation coefficient (r) values and the dynamic range covered in the plot clearly indicate the benefit of using PIPs for radiometric inter-calibration. The r values for all bands are in excess of 0.99, and the uncertainties (one-sigma error) of regression slopes are all less than 1%. The regression results demonstrate that PIPs can be well used for inter-calibration and can accurately measure the calibration difference of the two sensors.

The ROI in this study has a large spatial range, about 2600km horizontally and 1600km vertically. Due to the constraint of VZA in the IR-MAD procedure, the spatial distribution of PIPs in a single day is limited and cannot be spread over the entire ROI. Given the nominal revisit cycle of FY-3A/B is 5.5 days, assuming that the sensor radiometric calibration is stable during the period, it is reasonable to aggregate PIPs from 5 consecutive days to further enhance the abundance of targets and dynamic range. Figure 5 presents two examples of the spatial distribution of aggregated PIPs at early and later stage of FY-3B's lifecycle. Although the orbit of the satellite has drifted during this period and the sensor has also experienced relatively large degradation, the spatial distribution of PIPs is consistent except in some cloud areas. This means that whether

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2024.3376580

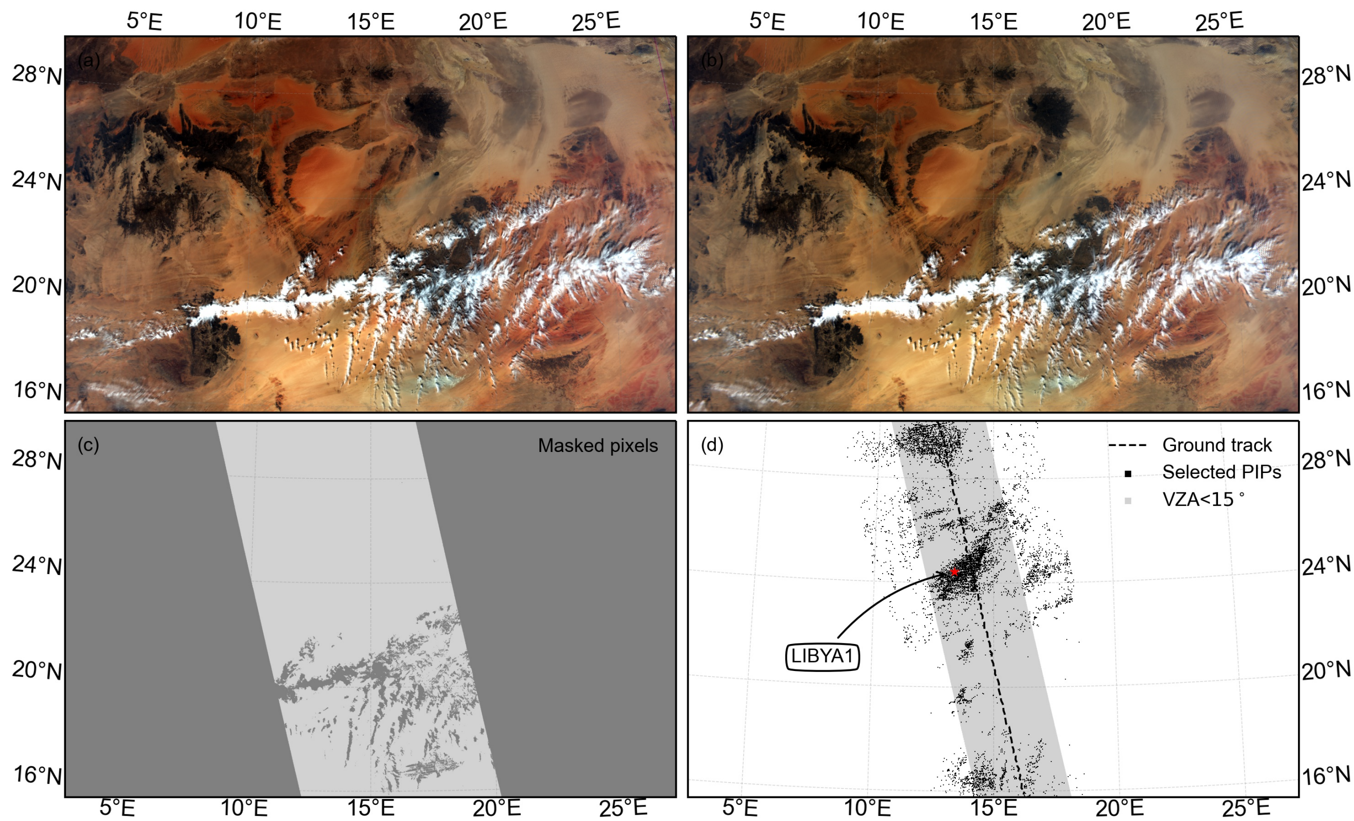

Fig. 3. True color images at the ROI from (a) FY-3B/MERSI and (b) FY-3B/VIRR on 20110123; (c) The mask used for statistic analysis of MAD method; (d) The spatial distribution of selected PIPs from IR-MAD method, only the pixels with VZA $< 15°$ are used for regression. PIPs: pseudo-invariant pixels; VZA: view zenith angle.

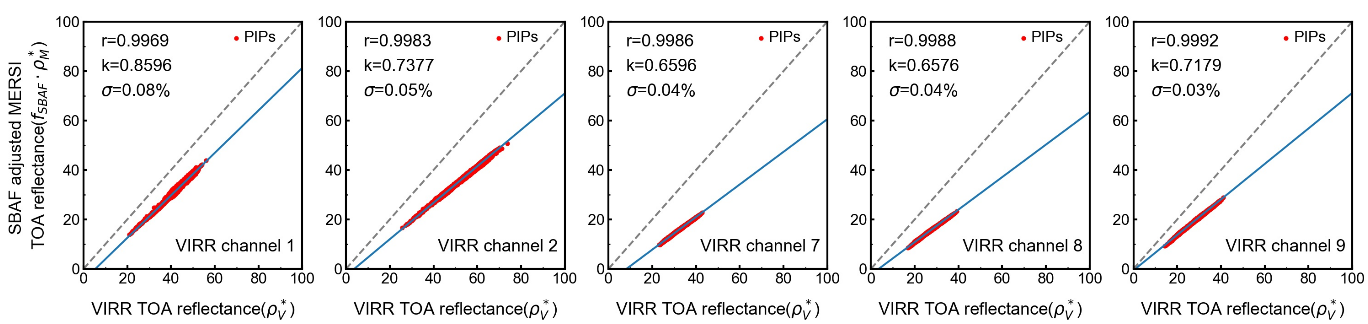

Fig. 4. Regression results of the PIPs selected by IR-MAD on 20110123. The number of PIPs is 7387. PIPs: pseudo-invariant pixels; r: correlation coefficient; k: regression slope; σ: uncertainty (one-sigma error) of k.

a target is selected as PIPs depends on the intrinsic properties of the target, which has nothing to do with the radiometric performance of the sensors.

Figure 6 shows the regression results of the aggregated PIPs from 5 consecutive days. In contrast to single-day PIPs for regression in Figure 4, the 5-day aggregated PIPs do not deviate from each other and still converge to the same linear relationship. What is more, the regression results of aggregated PIPs have wider reflectance dynamic range for all bands, and the uncertainty of the regression slope is no more than 0.4%. These superiorities indicate that it is reasonable and necessary to aggregate multi-day PIPs for regression analysis, especially for long-term inter-calibration with time span of several years.

B. IR-MAD vs. SNO-x

The most related work to our study is the SNO-x method, which extends the SNO analysis to the low latitude desert and ocean sites, and sets a number of criteria to choose spatially uniform ROIs to evaluate the bias between two sensors [8]. In this study, based on the potential linear relationship between two sensors, the IR-MAD technique was used to statistically selects PIPs to inter-calibrate the sensor. Here we compare the performance of the IR-MAD based method and the SNO-x based method for inter-calibration. There are multiple criteria

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2024.3376580

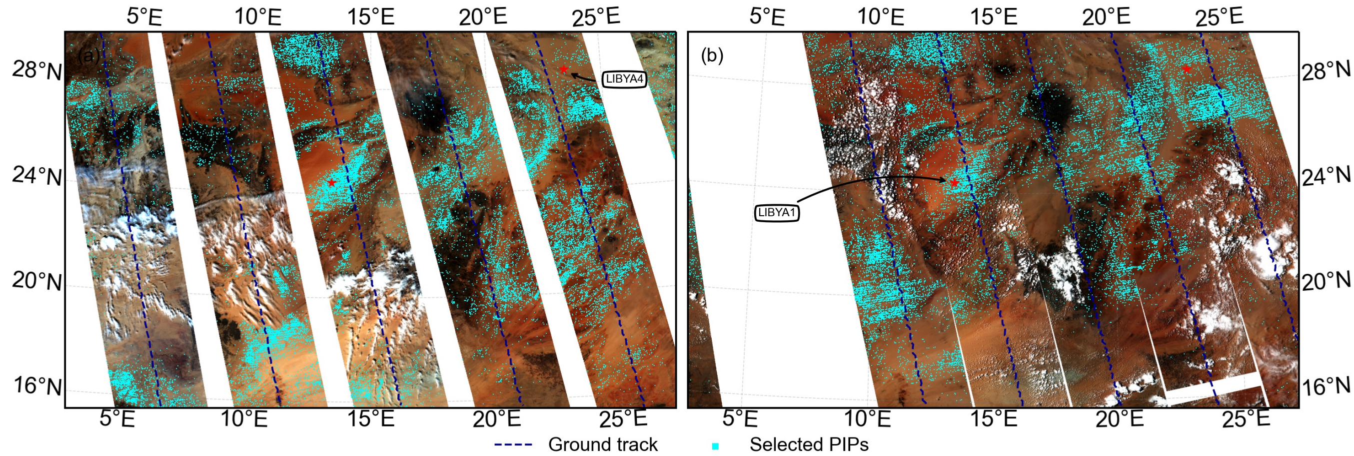

Fig. 5. The aggregated PIPs selected by IR-MAD overlay the concurrent true color images (VZA $< 15°$) of FY-3B/MERSI from five consecutive days: (a) 20110121-20110125 at early stage of FY-3B's lifecycle; (b) 20180728-20180801 in the later of FY-3B's lifecycle.

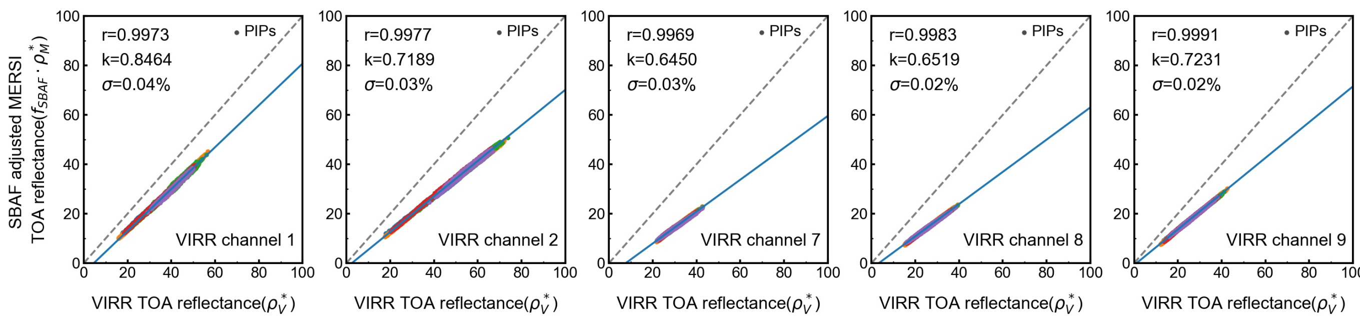

Fig. 6. Regression results of the aggregated PIPs selected by IR-MAD from 5 days (20110121, 20110122, 20110123, 20110124, 20110125). Different colored dots represent PIPs from different days. The total number of PIPs is 29171.

established to select the ROIs for SNO-x analysis as outlined in the original paper [8], including: 1) Scan angle difference, time difference and colsest-matching distance limit for geospatial matching. These are inherently satisfied as the two sensors used for intercomparison are onboard the same satellite; 2) The size of the ROI is 9 km×9 km, the solar zenith is less than $80°$ and the spatial uniformity should be less than 2%. These criteria align with those in the original paper; 3) The VZA is restricted to within $15°$ to ensure a fair comparison with the IR-MAD method. We set the moving stride to 5 (which means there is overlap) to increase the number of regression samples. The mean reflectance for each ROI is calculated for regression.

Figure 7 shows a comparison of the spatial distribution of matched pixels selected by two methods on January 21, 2011. The main reason for choosing this case is that it covers diverse land cover types, rather than a homogeneous desert area, which better reflects the superiority of our method. Figure 7 illustrates that in scenes containing complex samples, the SNO-x method selects fewer spatially uniform ROIs and the sample type is relatively small, mainly comprising desert targets. In contrast, the IR-MAD method can automatically select appropriate pseudo-invariant pixel-level samples for regression without requiring spatial uniformity, resulting in a greater number of samples with a wider range of types.

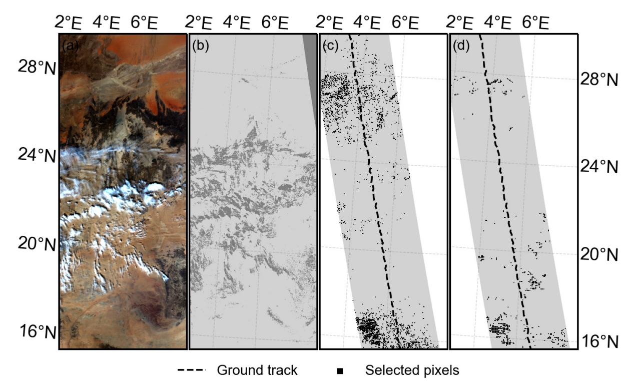

Fig. 7. A comparison of the spatial distribution of pixels selected by the SNO-x and IR-MAD methods on 20110121. (a) True color image from FY-3B/MERSI; (b) The mask used before selection; (c) Pixels selected by IR-MAD method. (d) Pixels selected by SNO-x method. The gray area in (c) and (d) represents the area where the VZA $< 15°$.

Figure 8 presents the regression results of the selected samples from one day by the two methods. It can be seen intuitively that the TOA reflectance of the regression samples selected by the IR-MAD method has a wider dynamic range, indicating a greater diversity in the selected samples.

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2024.3376580

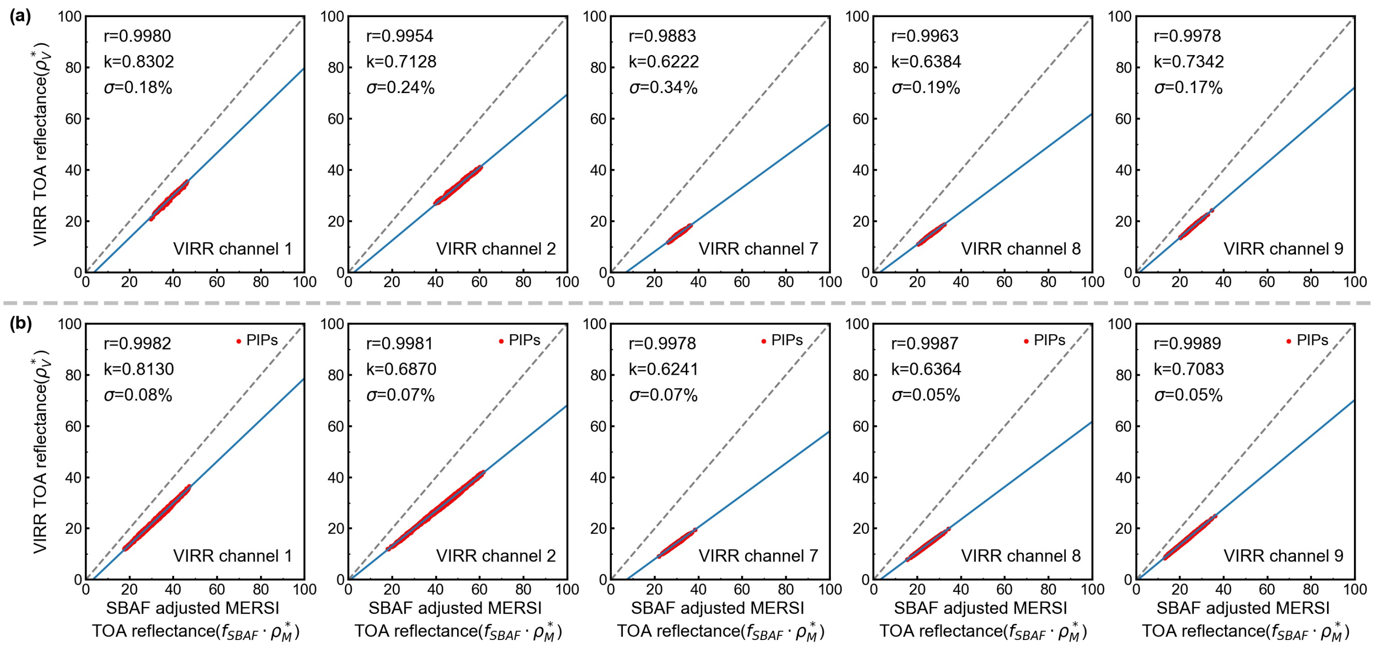

Fig. 8. Comparison of the regression results for FY-3B on 20110121, where the samples are selected by the two different methods: (a) SNO-x, with 812 samples; (b) IR-MAD method, with 3175 samples.

TABLE III
COMPARISON OF REGRESSION RESULTS OBTAINED BY APPLYING THE SNO-X METHOD AND THE IR-MAD METHOD FOR FY-3B OVER FIVE CONSECUTIVE DAYS.

date	method	number of samples	channel 1		channel 2		channel 7		channel 8		channel 9	
			σ(%)	range [min, max]	σ(%)	range [min, max]	σ(%)	range [min, max]	σ(%)	range [min, max]	σ(%)	range [min, max]
20110121	SNOX-x	812	0.18	[29.6, 46.2]	0.24	[39.6, 60.5]	0.34	[26.1, 36.3]	0.19	[20.4, 32.2]	0.17	[20.1, 34.6]
	IR-MAD	3875	0.08	[17.6, 47.3]	0.07	[17.9, 61.7]	0.07	[21.9, 38.5]	0.05	[15.5, 34.3]	0.05	[13.1, 36.0]
20110122	SNOX-x	3007	0.1	[32.9, 56.9]	0.05	[41.8, 72.6]	0.06	[26.7, 42.5]	0.03	[21.1, 39.5]	0.04	[20.9, 42.8]
	IR-MAD	4307	0.05	[15.9, 56.6]	0.02	[17.6, 72.1]	0.04	[21.7, 42.6]	0.03	[15.1, 39.5]	0.04	[12.0, 42.6]
20110123	SNOX-x	3436	0.21	[30.4, 57.9]	0.12	[42.7, 75.7]	0.06	[25.2, 41.5]	0.05	[19.9, 37.9]	0.05	[19.4, 40.7]
	IR-MAD	7387	0.08	[20.8, 56.1]	0.05	[25.7, 73.7]	0.04	[23.2, 42.8]	0.04	[16.8, 39.2]	0.03	[14.3, 41.1]
20110124	SNOX-x	2822	0.12	[31.2, 50.6]	0.11	[40.7, 66.1]	0.08	[25.9, 42.4]	0.07	[20.6, 36.7]	0.08	[21.4, 37.8]
	IR-MAD	5203	0.08	[18.3, 50.6]	0.05	[20.3, 66.4]	0.07	[22.3, 40.9]	0.05	[16.0, 36.7]	0.05	[13.6, 38.5]
20110125	SNOX-x	5991	0.13	[32.4, 50.1]	0.2	[42.1, 66.8]	0.06	[24.8, 42.2]	0.06	[19.5, 37.9]	0.05	[20.3, 37.6]
	IR-MAD	8399	0.1	[29.0, 49.9]	0.09	[37.7, 66.9]	0.04	[21.9, 42.7]	0.04	[15.9, 38.3]	0.04	[15.9, 37.9]

σ: uncertainty (one-sigma error) of regression slope.
range: TOA reflectance range of regression samples.

Furthermore, while the SNO-x method selects samples with overlapping, the number of samples is fewer than with IR-MAD, particularly in more complex scenes, e.g. those with higher cloud coverage where it is challenging to obtain large spatially uniform regions. With regards to regression results, the IR-MAD method has a very consistent regression slope with the SNO-x method, and the uncertainty of the regression slope is smaller, which also confirms the correctness of the samples selected by the IR-MAD method.

Table III presents the comparison of regression results obtained by applying the SNO-x method and the IR-MAD method from five consecutive days. It is evident that the IR-MAD method has significant advantages over the SNO-x method in terms of the number of regression samples, the dynamic range of TOA reflectance for each channel, and the uncertainty of the regression slope. Moreover, this method requires no prior knowledge of the surface and is globally applicable.

V. RESULTS

A. Long-term Time Series Results of FY-3A&3B VIRR

By performing IR-MAD procedure on all image pairs and subsequently aggregate 5-day PIPs for regression, a large database of $m_i(t)$ can be created. Note that the MERSI nominal TOA reflectance (ρ_M^*) has been compensated by f_{SBAF} before, thus the $m_i(t)$ here represents the measurement of $\frac{f_0}{S_{Vi}(t)}$. For FY-3A data with a time span of 6 years there are 439 5-day comparisons, and for FY-3B data with a time span of 8 years there are a total of 564 5-day comparisons. Analysis of these data can provide long-term inter-calibration results of the two sensors.

The long-term time series of $m_i(t)$ for FY-3A and FY-3B are shown in Figure 9. The gaps in the curves are due to missing data. It can be seen that there is a decreasing trend

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2024.3376580

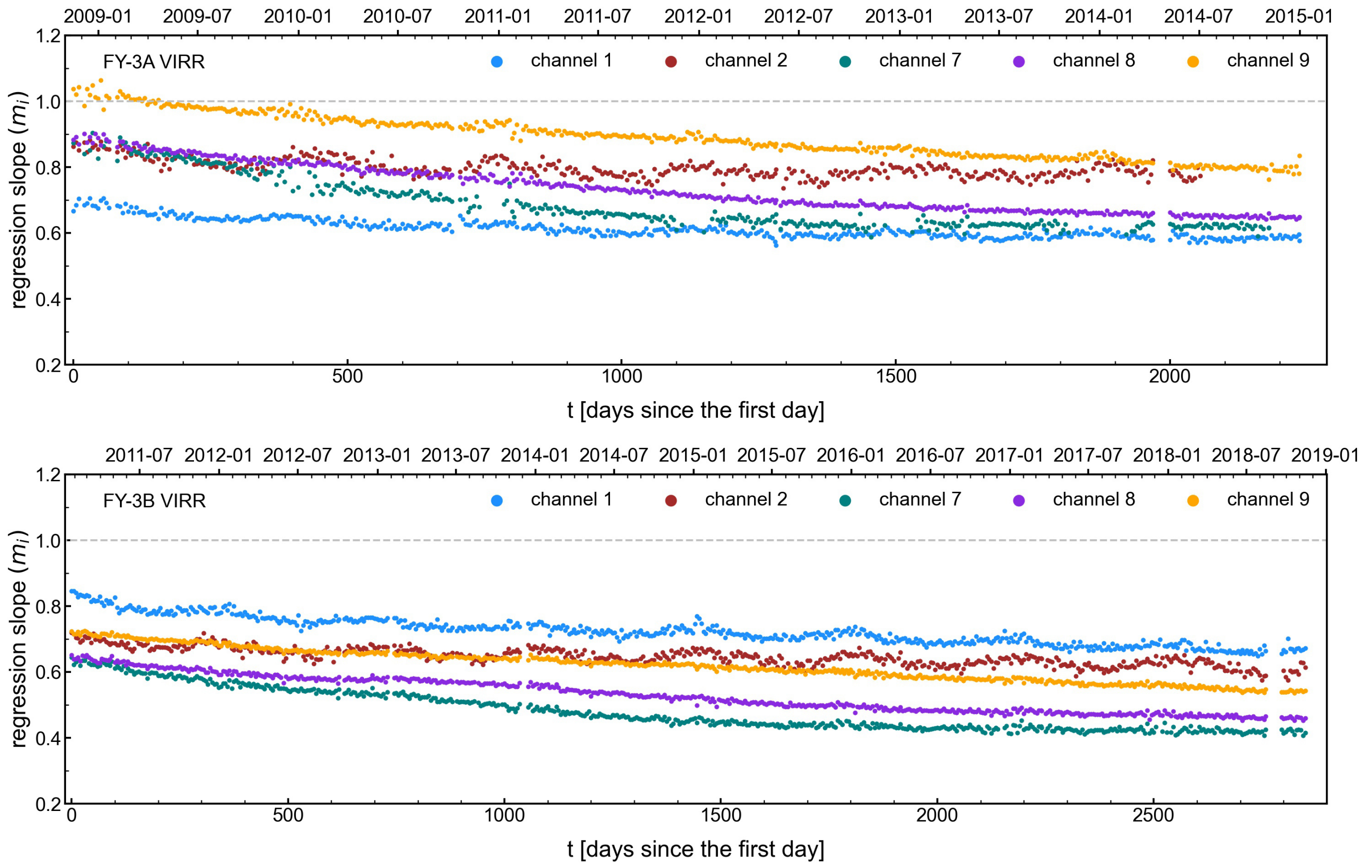

Fig. 9. The long-term time series of regression slope ($m_i(t)$) for FY-3A and FY-3B, which is a measurement of the relative calibration factor between MERSI and VIRR and the sensor degradation of VIRR itself.

in all bands for both sensors. This is because that all bands of VIRR have different degrees of degradation. In addition to long-term trends, periodic seasonal oscillations exist in some bands, especially for channel 1 and 2. Examination of the fitting error and potential overfitting problem shows that a polynomial of order 4 can best capture the trending pattern. For the regression intercepts in each comparison, they are attributed to the unaccounted for change in the zero-radiance response of VIRR after launch, which should generally remain stable over time in orbit. However, due to the degradation of VIRR itself, there would be a trend in the regression intercepts over time (not shown). In this study, we employed a second-order polynomial fitting to capture the trend of the regression intercept for correction. The coefficients of the fitted polynomials and bias statistics can be seen in Table IV.

As mentioned above, the $m_i(t)$ is a measurement of the relative calibration factor (f_0) and the sensor degradation ($1/S_{Vi0}$). The VIRR degradation can be obtained independently by the approach of [36]and [30]. Therefore, we can obtain the relative calibration factor (f_0) by dividing $m_i(t)$ by sensor degradation ($1/S_{Vi0}$). Figure 10 shows the long-term time series of f_0. Since we used fixed calibration coefficients at early operational stage to calculate the nominal TOA reflectance, f_0 represents the relative calibration factor of the two sensors at that time, which expects to be a constant. Due to the limitations of the polynomial functions used to fit the $m_i(t)$ and sensor degradation ($1/S_{Vi0}$) trending, they can not fully capture the seasonal variation. Consequently, periodic seasonal oscillations may manifest in the time series of the f_0. We tried to fit the trending of f_0 with polynomials of different orders, all tend to be a constant (as expected), which demonstrate the validity of the time series of $m_i(t)$ we derived.

Furthermore, we can utilize the estimated f_0 value to calculate the VIRR degradation and compare it with our previous work [30]. Unlike the IR-MAD technique applied in this study for inter-calibration, the IR-MAD technique is proposed to select temporally invariant pixels from the bitemporal satellite images by the same sensor to assess the sensor response degradation during the acquisition time interval [30]. By dividing the polynomials of $m_i(t)$ by the estimated f_0, we can obtain the sensor degradation as a function of time. The outcomes are depicted in Figure 11, where the solid lines correspond to the degradation calculated in this study and the dashed lines represent the outcomes obtained using the method of [30]. The highly consistent VIRR degradation curves obtained from the two methods further confirm the effectiveness of this approach.

Table V presents a quantitative comparison of the degradation results of VIRR obtained in this study with those from the other two different methods, including a multi-site calibration method [41], [42] and a IR-MAD based method for sensor degradation tracking [30], and their results have been unified to the same time range. It can be seen that

This article has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2024.3376580

TABLE IV
COEFFICIENTS OF THE FITTED POLYNOMIALS FOR LONG-TERM TIME SERIES OF $m_i(t)$ AND REGRESSION OFFSET, ALONG WITH THE FITTING BIAS STATISTICS FOR $m_i(t)$.

VIRR band	$m(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4$					$offset(t) = b_0 + b_1 t + b_2 t^2$			bias mean (%)	bias std (%)
	a_0	$a_1(\times 10^{-4})$	$a_2(\times 10^{-7})$	$a_3(\times 10^{-11})$	$a_4(\times 10^{-14})$	b_0	$b_1(\times 10^{-4})$	$b_2(\times 10^{-7})$		
fy3a/ch1	0.6858	-1.5767	1.1953	-4.7703	0.7568	-2.2214	8.1813	-3.6872	0.0001	1.6394
fy3a/ch2	0.8554	-1.4257	1.1732	-5.2327	1.0301	-2.4014	16.315	-8.1611	0.0000	2.3243
fy3a/ch7	0.8856	-3.2868	0.5078	6.8015	-2.2115	-5.0486	28.606	-14.152	0.0001	1.8337
fy3a/ch8	0.8848	-1.8592	0.2292	0.9403	-0.1600	-7.9868	4.1352	-2.0237	0.0000	0.8153
fy3a/ch9	1.0283	-2.2065	1.4163	-6.5519	1.1270	-5.6542	21.537	-7.2687	0.0000	1.0514
fy3b/ch1	0.8261	-1.9008	1.5373	-6.3876	0.9259	-2.6697	-2.6514	-0.1468	0.0000	1.5873
fy3b/ch2	0.6977	-0.8124	5.0772	-1.6215	0.1562	-1.2840	3.5416	0.6355	0.0000	2.3130
fy3b/ch7	0.6264	-1.5049	-0.0252	2.3268	-0.4491	-4.4480	9.0079	-2.7429	-0.0005	1.5665
fy3b/ch8	0.6376	-1.0266	0.3000	-1.5524	0.3598	-2.0573	1.8634	-0.5125	0.0002	1.2841
fy3b/ch9	0.7232	-1.4534	0.9793	-4.2652	0.6441	-1.0563	-0.9416	-0.3341	0.0000	0.7523

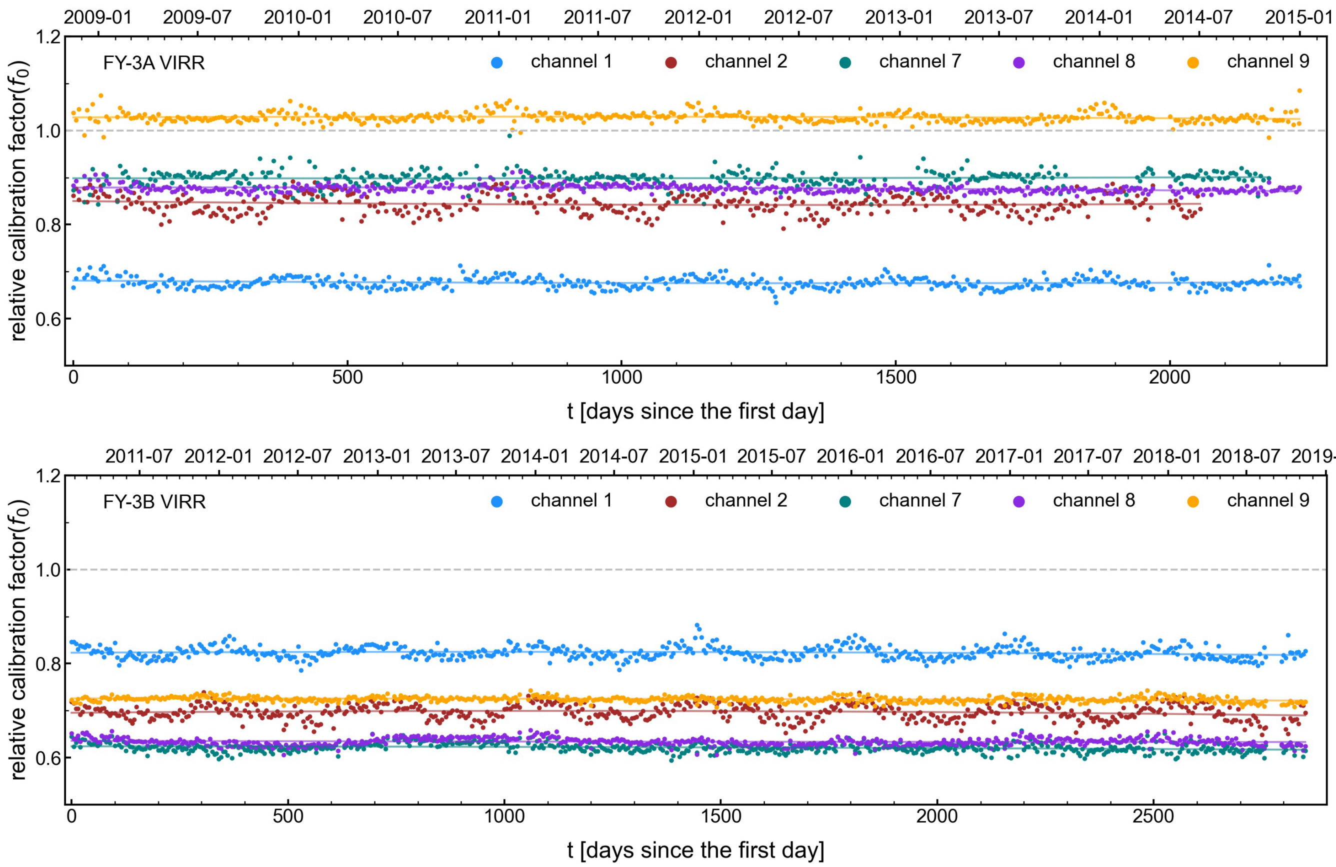

Fig. 10. The long-term time series of relative calibration factor (f_0) for FY-3A and FY-3B, which expects to be a constant. See main text for more explanation about f_0.

the degradation results of VIRR obtained in this study are very consistent with the other methods. Compared to the results of the other two methods, the annual degradation for most channels differs by less than 0.1%. For channels with significant seasonal fluctuations (such as channel 2), the difference can reach around 0.3%. As for channel 7, which is a short-wave channel, it is strongly affected by atmospheric absorption and scattering, resulting in a difference of around 0.5%. This verifies the consistency of the VIRR degradation results obtained by this method with other methods.

B. TOA reflectance trending validation over PICS

As an independent calibration method of on-orbit calibration, ground sites with suitable characteristics on Earth are often used to evaluate and validate the post-launch radiometric performance of satellite sensors. Monitoring the long-term time series of TOA reflectance at PICS is an effective approach to verify the radiometric calibration consistency between different sensors. In the ROI of this study, there are two PICS identified by the CEOS, namely LIBYA1 and LIBYA4. They are desert sites consisting of sand dunes and devoid of vegetation. These sites have been extensively studied and used as post-launch calibration sites for satellite optical sensors to evaluate the long-term stability and inter-comparisons. Due to

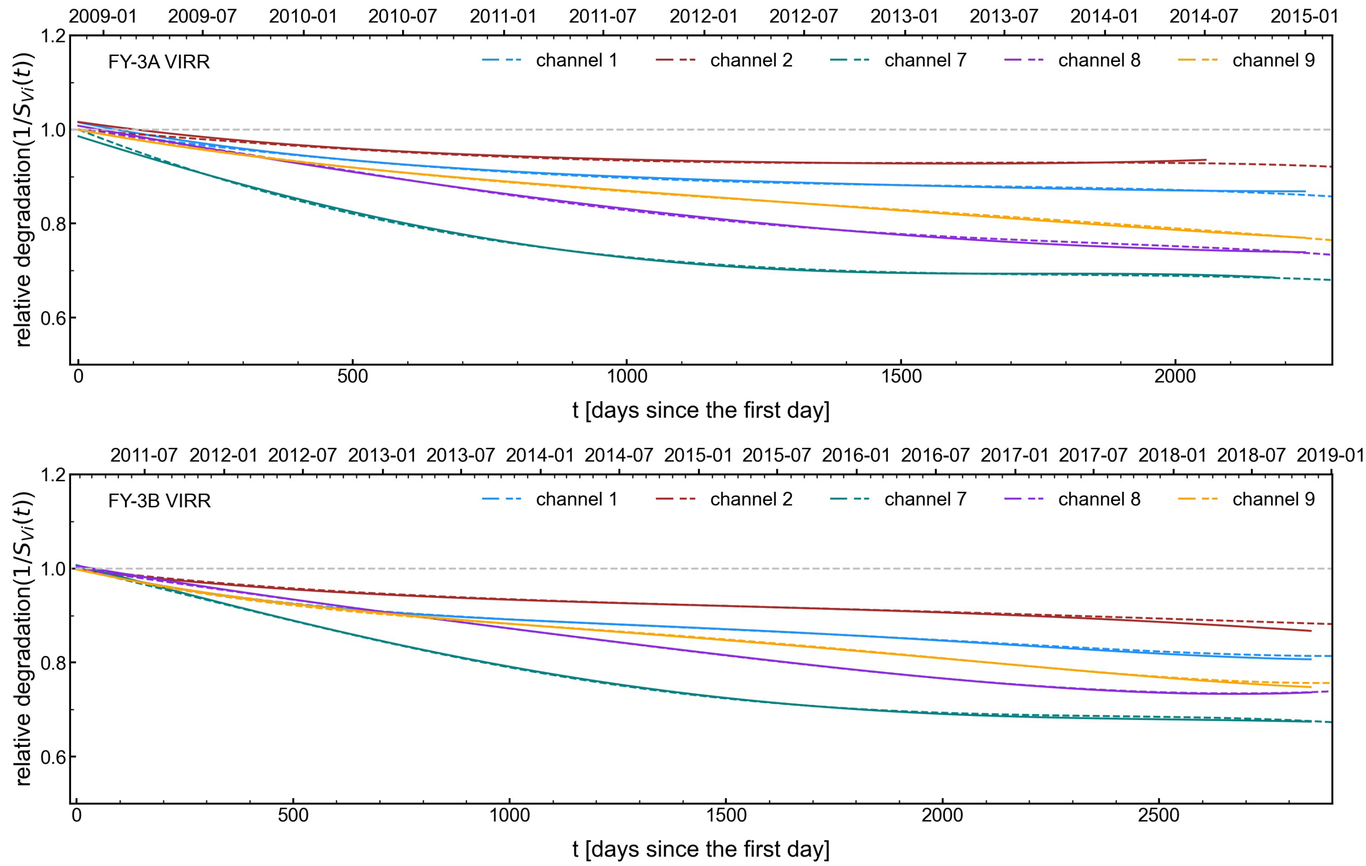

Fig. 11. The degradation curves over time of the VIRR on FY3A and FY3B. The solid lines correspond to the degradation obtained in this study and the dashed lines represent the results using the method of [30].

TABLE V
COMPARISON OF THE TOTAL AND ANNUAL DEGRADATION OF VIRR OBTAINED USING THREE DIFFERENT METHODS. FY3A: FROM NOVEMBER 16, 2008, TO DECEMBER 1, 2014. FY3B: FROM JANUARY 25, 2011, TO DECEMBER 1, 2017.

VIRR band	total degradation			annual degradation		
	method 1 ([41])	method 2 ([30])	this study	method 1 ([41])	method 2 ([30])	this study
fy3a/ch1	14.06%	13.74%	14.44%	2.33%	2.27%	2.39%
fy3a/ch2	9.48%	7.58%	7.17%	1.57%	1.25%	1.19%
fy3a/ch7	30.26%	31.66%	30.66%	5.01%	5.24%	5.07%
fy3a/ch8	27.11%	26.02%	26.62%	4.49%	4.31%	4.40%
fy3a/ch9	22.18%	22.81%	22.76%	3.67%	3.77%	3.77%
fy3b/ch1	18.08%	17.65%	18.26%	2.64%	2.57%	2.66%
fy3b/ch2	13.54%	10.64%	11.20%	1.98%	1.55%	1.63%
fy3b/ch7	35.67%	31.54%	32.54%	5.20%	4.60%	4.75%
fy3b/ch8	28.34%	26.30%	26.85%	4.13%	3.84%	3.92%
fy3b/ch9	23.07%	23.03%	22.99%	3.37%	3.36%	3.35%

space limitations, the long-term time series of TOA reflectance over Libya 4 for FY-3B are presented here. FY-3B was selected due to its longer time span, and LIBYA4 has been the most commonly used calibration site in recent years. Similar results can also be obtained for FY-3A and LIBYA1, which will be presented in the table.

By selecting clear-sky samples with a sensor zenith angle less than 20 degrees during satellite overpasses, Figure 12 shows the long-term time series of the TOA reflectance of FY-3B/VIRR over LIBAYA4 before and after inter-calibration, with the MERSI used as a reference. To mitigate seasonal oscillations arising from BRDF effect or orbital drift, the TOA reflectance has been characterized by the solar zenith angle and normalized to the TOA reflectance at a solar zenith angle of 30 degrees. From the figure, it can be observed that before inter-calibration, due to the degradation of the VIRR itself, there is a significant downward trend in each channel, while MERSI as the reference sensor has a very stable and flat long-term response. After inter-calibration, the trend of VIRR has been eliminated and it has a consistent radiometric response with MERSI, indicating that our inter-calibration method is effective. The second column of Figure 12 shows the relative deviation of TOA reflectance between MERSI and inter-calibrated VIRR, which exhibits a seasonal oscillations pattern. The pattern becomes more pronounced in the later stage of the satellite's life cycle, which may be related to the orbital drift of FY-3B in its later phase. Table VI presents the quantitative results of the relative deviation between MERSI and inter-calibrated VIRR on FY-3A and FY-3B over LIBIYA1 and LIBAYA4. It can be seen that for most channels, the mean deviation is less than 1%, with a standard deviation of less than 2%. For some channels with significant seasonal fluctuations, the deviation may be slightly larger, but the mean deviation is also less than 2%, with a standard deviation of less than 2.5%. The mechanism behind the significant seasonal fluctuations and amplitude of TOA reflectance bias will be further investigated. The quantitative results of the two satellites over the two PICS have also verified the accuracy and effectiveness of our method.

TABLE VI
QUANTITATIVE RESULTS OF THE RELATIVE DEVIATION BETWEEN MERSI AND INTER-CALIBRATED VIRR ON FY-3A AND FY-3B OVER LIBYA1 AND LIBYA4.

VIRR band	LIBYA1		LIBYA4	
	bias mean	bias std	bias mean	bias std
fy3a/ch1	-0.578%	0.866%	0.845%	0.888%
fy3a/ch2	0.725%	1.623%	-1.200%	1.626%
fy3a/ch7	0.667%	1.550%	-0.790%	1.297%
fy3a/ch8	1.009%	1.096%	-0.371%	0.915%
fy3a/ch9	0.767%	1.034%	-0.654%	0.872%
fy3b/ch1	-1.399%	0.855%	1.401%	0.826%
fy3b/ch2	0.479%	1.787%	-0.434%	1.795%
fy3b/ch7	1.480%	1.750%	0.640%	1.532%
fy3b/ch8	1.341%	1.360%	0.531%	1.319%
fy3b/ch9	0.557%	1.213%	-0.093%	1.112%

C. Uncertainty Analysis

In this inter-calibration, the uncertainty primarily originates from the radiometric calibration uncertainty of the reference sensor and the processing procedure of the inter-calibration. Here we mainly analyze the uncertainty introduced by several factors in our processing method, including: geometric misregistration, spectral band differences, atmospheric conditions and viewing geometry (or BRDF) effects during overpass. In this study, since the two sensors are onboard the same platform, the differences caused by atmospheric conditions and BRDF effects are negligible.

MERSI and VIRR have a little difference in spatial resolution, making it challenging to achieve perfect pixel matching between the two sensors, especially in off-nadir areas. To minimize the effects of different spatial resolutions, our method limits the near-nadir (VZA$\leq$15) pixels for analysis and regression. Furthermore, the IR-MAD technique statistically selects PIPs that satisfy the underlying linear relationship (Equal 9) in each channel, hence pixels with larger differences due to misregistration will not be selected as PIPs. As a result, in the presence of misregistration effect, IR-MAD tends to select PIPs from spatially uniform areas. Similarly, when atmospheric disturbances and BRDF effects are present, IR-MAD also tends to select pixels with the smallest possible impact of these effects. Thus, the uncertainty introduced by these effects is implicitly reduced during the PIPs selection of IR-MAD.

One of the primary sources of uncertainty in our method is the differences in relative spectral responses. Despite compensating with SBAF, the SBAFs for different land cover types are not the same. Hence, when the study area is vast and encompasses diverse land cover types, IR-MAD favors selecting pixels with flat spectral signature to minimize such effects.

The uncertainty arising from the above effects is reflected in the residual standard deviation of the regression results. For the 439 5-day comparisons of FY-3A and the 565 5-day comparisons of FY-3B, the mean uncertainty for the matching bands are listed on the first row of Table VII. For long-term inter-calibration results, additional uncertainties are introduced by polynomial fitting, mainly due to seasonal oscillations of the sensor, which are listed on the second row of Table VII. The overall uncertainties for the matching bands are listed on the last row of Table VII.

VI. DISCUSSION

Unlike other previous studies that used IR-MAD method to select temporally invariant pixels from bitemporal image of one sensor, this study employed the IR-MAD method to select PIPs from multispectral images acquired at the same time by two different sensors for inter-calibration. Starting from the formula of the inter-calibration problem, we demonstrate the reasonableness of employing IR-MAD technique to select PIPs for inter-calibration owing to the linear scale invariance property of MAD transformation. The PIPs represent pixels that conform to the potential linear relationship across all matching bands in inter-calibration, under the differences caused by a combination of relative spectral response characteristics of the two sensors, spectral signature of the target, atmospheric conditions, and viewing geometry.

The PIPs are selected statistically from the image pair without a priori knowledge of the target pixel, which are determined based on the existence of a certain underlying linear relationship between the matching bands of the two images. Therefore, for different pixels within an image pair, they may not conform to this linear relationship due to variations in surface characteristics or differences in atmospheric influence. As a result, the locations of PIPs may vary with different image pairs. It should be noted that when the two images in an image pair are strongly affected by atmospheric conditions or when there is a lack of a sufficient number of samples for statistical analysis due to extensive cloud cover, the number of selected PIPs will decrease, which means it is challenging for IR-MAD to find the underlying linear relationship between the two images. Using these PIPs for regression will result in greater uncertainty. Therefore, such image pairs will be excluded. To examine which areas and surface types are most frequently identified as PIPs, a density map showing the spatial frequency distribution of PIPs of all FY-3B image pairs is depicted in Figure 13. Similar results can also be obtained with FY-3A.

From Figure 13, it is evident that PIPs are distributed not only in bright target areas such as deserts but also encompass dark target types like volcanic surfaces. Therefore, inter-calibration based on PIPs offers a broader dynamic range.

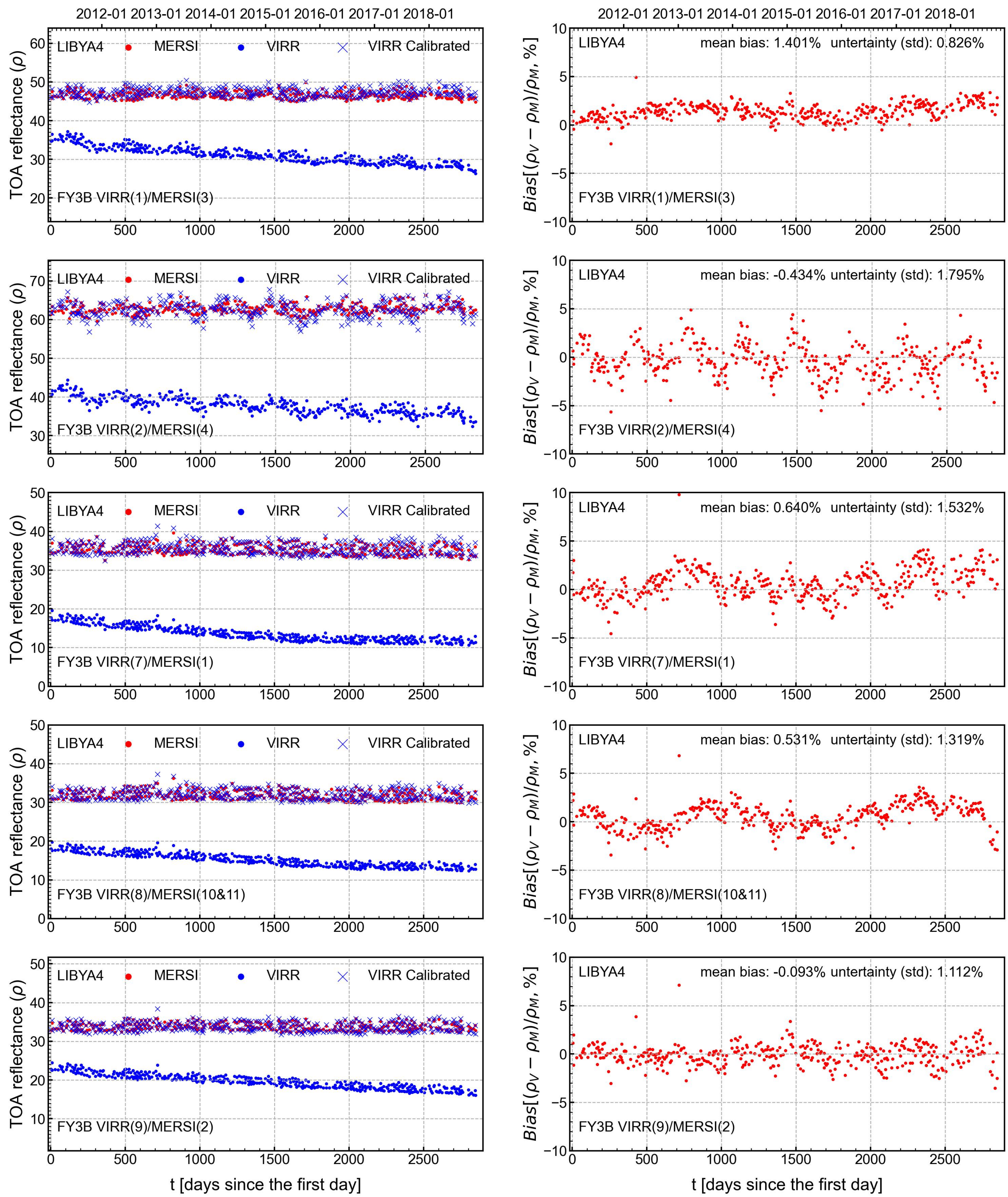

Fig. 12. The long-term time series of the TOA reflectance of FY-3B/VIRR over LIBAYA4 before and after inter-calibration, with the MERSI used as a reference. The second column shows the relative deviation of TOA reflectance between MERSI and inter-calibrated VIRR.

TABLE VII
UNCERTAINTY OF THE PROPOSED IR-MAD METHOD FOR INTER-CALIBRATION.

Uncertainty	Source	FY3A VIRR					FY3B VIRR				
		ch1(%)	ch2(%)	ch7(%)	ch8(%)	ch9(%)	ch1(%)	ch2(%)	ch7(%)	ch8(%)	ch9(%)
uncertainty for image pair	misregistration atmospheric conditions BRDF effect spectral band differences	1.46	1.30	0.69	0.68	0.65	1.50	1.23	0.85	0.78	0.89
uncertainty for long-term inter-calibration	polynomial fitting	1.64	2.32	1.83	0.81	1.05	1.59	2.31	1.57	1.28	0.75
	total error (root sum of squares)	2.20	2.67	1.96	1.06	1.24	2.19	2.62	1.78	1.50	1.17

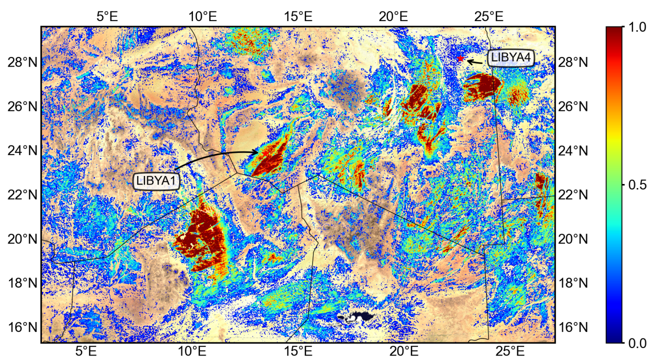

Fig. 13. The density map of the spatial frequency distribution of PIPs for all image pairs of FY-3B.

The primary hotspot regions are predominantly located within desert areas, which aligns with our expectations. Deserts exhibit flatter spectral profiles, minimizing the impact of SRFs differences between the two sensors. Regarding the spatial distribution pattern of PIPs, our method does not rely on large, spatially uniform regions. Instead, it selects samples at the pixel level, resulting in a higher sample quantity, a richer variety of land surface types, and a broader dynamic range compared to the SNOx method. Furthermore, in addition to internationally recognized PICS such as LIBYA1 and LIBYA4 being found within the distribution area of PIPs, there are several extensive PIPs hotspot regions. These areas hold the potential to serve as valuable calibration sites for future research.

Another advantage of this method is its ability to implicitly reduce uncertainty. Unlike SNO-x method, which relies on the spatial uniformity of the surface to select samples, IR-MAD selects samples based on the potential linear relationship between two images. Samples that no longer conform to this linear relationship due to effects such as misregistration, atmospheric disturbances, and BRDF are excluded from the PIPs selection. As a result, the uncertainty introduced by these effects is implicitly minimized during the PIPs selection process of IR-MAD. It's worth noting that this method employs a single SBAF for all PIPs samples, which is a major source of uncertainty in this method. However, as shown in Figure 13, IR-MAD tends to select targets like deserts that have relatively flat spectral profiles to reduce errors caused by the SRF differences. For long-term inter-calibration, variations in atmospheric conditions at different times are the primary source of uncertainty, as evident in Figure 9. For channels easily affected by atmospheric conditions, their regression slopes exhibit fluctuations and divergence (e.g. channels 1 and 2), while this phenomenon is less pronounced for channels that are less susceptible to atmospheric influence (channels 8 and 9).

This method is not limited to a specific geographical region. We also conducted research in the northwest region of China as our ROI and obtained similar results. The results can be found in the supplementary materials. Furthermore, this method is not restricted to any particular sensor or satellite platform. It can be applied to SNO and SNO-x events as well. The entire process is generic. The difference lies in the fact that, unlike the two sensors on the same platform in this study, when dealing with sensors on different platforms, additional considerations are required to account for the impact of BRDF effects. Therefore, in such cases, additional constraints should be added to minimize the influence of BRDF. This can be achieved by imposing constraints on the proximity of the two satellite orbits or by using samples from nadir observations for the analysis.

VII. CONCLUDING REMARKS

In this study, we propose a novel approach to the sensor radiometric inter-calibration based on PIPs using IR-MAD method. Due to the property of linear scale invariance, the IR-MAD method was proposed to statistically select PIPs for inter-calibration. The approach requires no prior knowledge of the inter-calibration targets. The PIPs do not depend on large spatially homogeneous areas and extends PICS to the pixel-level targets, resulting in a higher sample quantity, a richer variety of land surface types, and a broader dynamic range of reflectance. The PIPs are selected statistically and can implicitly reduce uncertainty of inter-calibration. This method is generic, not limited to any particular sensor, satellite platform or geographical region, which is particularly suitable for operational long-term inter-calibration.

The implementation on FY-3A&3B VIRR for inter-calibration (with the MERSI onboard the same platform as the reference) demonstrates the efficacy of our method. The results show that two widely used PICS for calibration, LIBYA1 and LIBYA4, have been automatically included in the PIPs. Moreover, despite the degradation of the sensors and drift of the satellite orbit over time, the spatial distribution of the PIPs remains consistent. The long-term time series of TOA reflectance over LIBYA1 and LIBYA4 shows that the inter-calibrated VIRR is in good agreement with MERSI, with a mean bias of less than 1% and an uncertainty of less than 2% for most channels. For channels with significant seasonal oscillations, the uncertainty is also less than 2.5%. Further exploration of the underlying mechanisms for the seasonal fluctuations in the long-term inter-calibration results is still needed. It should be reiterated that although this study used two sensors onboard the same platform for inter-calibration, our method is also applicable to other situations where similar sensors for inter-calibration are onboard different platforms, such as SNO or SNO-x events.

Data Availability

The data used in this study are all publicly available in the FY-3A&B satellite archive at FENGYUN Satellite Data Center. These datasets were derived from the public domain resources: http://satellite.nsmc.org.cn/portalsite/default.aspx?currentculture=en-US.

References

[1] P. N. Slater, S. F. Biggar, J. M. Palmer, and K. J. Thome, "Unified approach to absolute radiometric calibration in the solar-reflective range," *Remote Sensing of Environment*, vol. 77, no. 3, pp. 293–303, 2001.

[2] P. Teillet, J. Barker, B. Markham, R. Irish, G. Fedosejevs, and J. Storey, "Radiometric cross-calibration of the landsat-7 etm+ and landsat-5 tm sensors based on tandem data sets," *Remote sensing of Environment*, vol. 78, no. 1-2, pp. 39–54, 2001.

[3] A. K. Heidinger, C. Cao, and J. T. Sullivan, "Using moderate resolution imaging spectrometer (modis) to calibrate advanced very high resolution radiometer reflectance channels," *Journal of Geophysical Research: Atmospheres*, vol. 107, no. D23, pp. AAC 11–1–AAC 11–10, 2002.

[4] C. Cao, H. Xu, J. Sullivan, L. McMillin, P. Ciren, and Y.-T. Hou, "Intersatellite radiance biases for the high-resolution infrared radiation sounders (hirs) on board noaa-15,-16, and-17 from simultaneous nadir observations," *Journal of Atmospheric and Oceanic Technology*, vol. 22, no. 4, pp. 381–395, 2005.

[5] X. Wu and F. Sun, "Post-launch calibration of GOES Imager visible channel using MODIS data," in *Earth Observing Systems X*, J. J. Butler, Ed., vol. 5882, International Society for Optics and Photonics. SPIE, 2005, p. 58820N.

[6] A. Wu, C. Cao, and X. Xiong, "Using MODIS to track calibration stability of the AVHRR on NOAA 15-18," in *Remote Sensing and Modeling of Ecosystems for Sustainability III*, W. Gao and S. L. Ustin, Eds., vol. 6298, International Society for Optics and Photonics. SPIE, 2006, p. 629812.

[7] A. K. Heidinger, W. C. Straka III, C. C. Molling, J. T. Sullivan, and X. Wu, "Deriving an inter-sensor consistent calibration for the avhrr solar reflectance data record," *International Journal of Remote Sensing*, vol. 31, no. 24, pp. 6493–6517, 2010.

[8] S. Uprety, C. Cao, X. Xiong, S. Blonski, A. Wu, and X. Shao, "Radiometric intercomparison between suomi-npp viirs and aqua modis reflective solar bands using simultaneous nadir overpass in the low latitudes," *Journal of Atmospheric and Oceanic Technology*, vol. 30, no. 12, pp. 2720–2736, 2013.

[9] P. Teillet, P. Slater, Y. Ding, R. Santer, R. Jackson, and M. Moran, "Three methods for the absolute calibration of the noaa avhrr sensors in-flight," *Remote sensing of Environment*, vol. 31, no. 2, pp. 105–120, 1990.

[10] P. N. Slater, S. F. Biggar, K. J. Thome, D. I. Gellman, and P. R. Spyak, "Vicarious radiometric calibrations of eos sensors," *Journal of Atmospheric and Oceanic Technology*, vol. 13, no. 2, pp. 349–359, 1996.

[11] P. Teillet, G. Fedosejevs, R. Gauthier, N. O'neill, K. J. Thome, S. F. Biggar, H. Ripley, and A. Meygret, "A generalized approach to the vicarious calibration of multiple earth observation sensors using hyperspectral data," *Remote Sensing of Environment*, vol. 77, no. 3, pp. 304–327, 2001.

[12] K. Thome, "Absolute radiometric calibration of landsat 7 etm+ using the reflectance-based method," *Remote Sensing of Environment*, vol. 78, no. 1-2, pp. 27–38, 2001.

[13] P. N. Slater, S. F. Biggar, J. M. Palmer, and K. J. Thome, "Unified approach to absolute radiometric calibration in the solar-reflective range," *Remote Sensing of Environment*, vol. 77, no. 3, pp. 293–303, 2001.

[14] C. J. Bruegge, N. L. Chrien, R. R. Ando, D. J. Diner, W. A. Abdou, M. C. Helmlinger, S. H. Pilorz, and K. J. Thome, "Early validation of the multi-angle imaging spectroradiometer (misr) radiometric scale," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 40, no. 7, pp. 1477–1492, 2002.

[15] K. J. Thome, D. L. Helder, D. Aaron, and J. D. Dewald, "Landsat-5 tm and landsat-7 etm+ absolute radiometric calibration using the reflectance-based method," *IEEE Transactions on geoscience and remote sensing*, vol. 42, no. 12, pp. 2777–2785, 2004.

[16] H. Cosnefroy, M. Leroy, and X. Briottet, "Selection and characterization of saharan and arabian desert sites for the calibration of optical satellite sensors," *Remote Sensing of Environment*, vol. 58, no. 1, pp. 101–114, 1996.

[17] A. Angal, X. Xiong, T. Choi, G. Chander, and A. Wu, "Using the sonoran and libyan desert test sites to monitor the temporal stability of reflective solar bands for landsat 7 enhanced thematic mapper plus and terra moderate resolution imaging spectroradiometer sensors," *Journal of Applied Remote Sensing*, vol. 4, no. 1, p. 043525, 2010.

[18] G. Chander, X. J. Xiong, T. J. Choi, and A. Angal, "Monitoring on-orbit calibration stability of the terra modis and landsat 7 etm+ sensors using pseudo-invariant test sites," *Remote Sensing of Environment*, vol. 114, no. 4, pp. 925–939, 2010.

[19] P. Minnis, D. R. Doelling, L. Nguyen, W. F. Miller, and V. Chakrapani, "Assessment of the visible channel calibrations of the virs on trmm and modis on aqua and terra," *Journal of Atmospheric and Oceanic Technology*, vol. 25, no. 3, pp. 385–400, 2008.

[20] Y. Kaufman and B. Holben, "Calibration of the avhrr visible and near-ir bands by atmospheric scattering, ocean glint and desert reflection," *International Journal of Remote Sensing*, vol. 14, no. 1, pp. 21–52, 1993.

[21] O. Hagolle, J.-M. Nicolas, B. Fougnie, F. Cabot, and P. Henry, "Absolute calibration of vegetation derived from an interband method based on the sun glint over ocean," *IEEE transactions on geoscience and remote sensing*, vol. 42, no. 7, pp. 1472–1481, 2004.

[22] X. Xiong, J. Sun, and W. Barnes, "Intercomparison of on-orbit calibration consistency between terra and aqua modis reflective solar bands using the moon," *IEEE Geoscience and remote sensing letters*, vol. 5, no. 4, pp. 778–782, 2008.

[23] T. C. Stone, W. B. Rossow, J. Ferrier, and L. M. Hinkelman, "Evaluation of isccp multisatellite radiance calibration for geostationary imager visible channels using the moon," *IEEE transactions on geoscience and remote sensing*, vol. 51, no. 3, pp. 1255–1266, 2013.

[24] C. Cao and A. K. Heidinger, "Intercomparison of the longwave infrared channels of modis and avhrr/noaa-16 using simultaneous nadir observations at orbit intersections," in *Earth Observing Systems VII*, vol. 4814. SPIE, 2002, pp. 306–316.

[25] C. Cao, X. Xiong, A. Wu, and X. Wu, "Assessing the consistency of avhrr and modis l1b reflectance for generating fundamental climate data records," *Journal of Geophysical Research: Atmospheres*, vol. 113, no. D9, 2008.

[26] A. K. Heidinger, W. C. Straka III, C. C. Molling, J. T. Sullivan, and X. Wu, "Deriving an inter-sensor consistent calibration for the avhrr solar reflectance data record," *International Journal of Remote Sensing*, vol. 31, no. 24, pp. 6493–6517, 2010.

[27] C. Dong, J. Yang, W. Zhang, Z. Yang, N. Lu, J. Shi, P. Zhang, Y. Liu, and B. Cai, "An overview of a new chinese weather satellite fy-3a," *Bulletin of the American Meteorological Society*, vol. 90, no. 10, pp. 1531–1544, 2009.

[28] L. Sun, M.-H. Guo, N. Xu, L.-J. Zhang, J.-J. Liu, X.-Q. Hu, Y. Li, Z.-G. Rong, Z.-H. Zhao *et al.*, "On-orbit response variation analysis of fy-3 mersi reflective solar bands based on dunhuang site calibration," *Spectroscopy and Spectral Analysis*, vol. 32, no. 7, pp. 1869–1877, 2012.

[29] X. Hu, L. Sun, J. Liu, L. Ding, X. Wang, Y. Li, Y. Zhang, N. Xu, and L. Chen, "Calibration for the solar reflective bands of medium resolution spectral imager onboard fy-3a," *IEEE Transactions on Geoscience and Remote sensing*, vol. 50, no. 12, pp. 4915–4928, 2012.

[30] J. Wang, X. Hu, Y. He, and K. Gao, "Response degradation analysis of fengyun-3a medium-resolution spectral imager based on intelligent detection of invariant pixels," *Acta Optica Sinica*, vol. 39, no. 9, p. 0912001, 2019.

[31] G. Chander, B. L. Markham, and D. L. Helder, "Summary of current radiometric calibration coefficients for landsat mss, tm, etm+, and eo-1 ali sensors," *Remote sensing of environment*, vol. 113, no. 5, pp. 893–903, 2009.

[32] G. Thuillier, M. Hersé, D. Labs, T. Foujols, W. Peetermans, D. Gillotay, P. Simon, and H. Mandel, "The solar spectral irradiance from 200 to

2400 nm as measured by the solspec spectrometer from the atlas and eureca missions," *Solar Physics*, vol. 214, pp. 1–22, 2003.

[33] G. Chander, N. Mishra, D. L. Helder, D. B. Aaron, A. Angal, T. Choi, X. Xiong, and D. R. Doelling, "Applications of spectral band adjustment factors (sbaf) for cross-calibration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 3, pp. 1267–1281, 2012.

[34] A. A. Nielsen, K. Conradsen, and J. J. Simpson, "Multivariate alteration detection (mad) and maf postprocessing in multispectral, bitemporal image data: New approaches to change detection studies," *Remote Sensing of Environment*, vol. 64, no. 1, pp. 1–19, 1998.

[35] M. J. Canty, A. A. Nielsen, and M. Schmidt, "Automatic radiometric normalization of multitemporal satellite imagery," *Remote sensing of Environment*, vol. 91, no. 3-4, pp. 441–451, 2004.

[36] M. Schmidt, E. A. King, and T. R. McVicar, "A method for operational calibration of avhrr reflective time series data," *Remote Sensing of Environment*, vol. 112, no. 3, pp. 1117–1129, 2008.

[37] A. A. Nielsen, "The regularized iteratively reweighted mad method for change detection in multi-and hyperspectral data," *IEEE Transactions on Image processing*, vol. 16, no. 2, pp. 463–478, 2007.

[38] M. J. Canty and A. A. Nielsen, "Automatic radiometric normalization of multitemporal satellite imagery with the iteratively re-weighted mad transformation," *Remote sensing of environment*, vol. 112, no. 3, pp. 1025–1036, 2008.

[39] X. Hu, L. Wang, J. Wang, L. He, L. Chen, N. Xu, B. Tao, L. Zhang, P. Zhang, and N. Lu, "Preliminary selection and characterization of pseudo-invariant calibration sites in northwest china," *Remote Sensing*, vol. 12, no. 16, p. 2517, 2020.

[40] H. Hotelling, "Relations between two sets of variates," in *Breakthroughs in statistics*. Springer, 1992, pp. 162–190.

[41] L. Wang, X. Hu, L. Chen, and L. He, "Consistent calibration of virr reflective solar channels onboard fy-3a, fy-3b, and fy-3c using a multisite calibration method," *Remote Sensing*, vol. 10, no. 9, 2018. [Online]. Available: https://www.mdpi.com/2072-4292/10/9/1336

[42] L. Sun, H. Qiu, R. Wu, J. Wang, L. Zhang, and P. Zhang, "Long-term consistent recalibration of virr solar reflectance data record for fengyun polar-orbiting satellites," *Journal of Meteorological Research*, vol. 35, pp. 926–942, 2021.