Radiometric Intercalibration Methodologies for High-Resolution Satellite Imagery in Precision Agriculture

Christian Massimiliano Baldin*, Vittorio Marco Casella

Department of Civil Engineering and Architecture, University of Pavia - 27100 Pavia, Italy
(christianmassimiliano.baldin, vittorio.casella)@unipv.it

Keywords: Radiometry, Calibration, Sentinel-2, PlanetScope, NDRE, Moran's I.

Abstract

This paper examines how to align PlanetScope and Sentinel-2 vegetation indices, focusing on the Normalized Difference Red Edge (NDRE) index, which is commonly used in precision agriculture for prescription maps. While Sentinel-2 is popular for crop monitoring, its low spatial resolution limits use in small or irregular fields. PlanetScope provides higher-resolution, more frequent imagery, but its sensor differs from the Sentinel-2, limiting compatibility with current research and tools. By testing three adjustment methods, the study shows that it is possible to align PlanetScope NDRE values with Sentinel-2: M1 (Linear Regression + Histogram Shifting + Histogram Matching), M2 (Histogram Matching), and M3 (per-band linear regression before index calculation). Two dates from 2022 were selected as representative seasonal extremes from the broader 2021–2023 dataset of 56 image pairs (Baldin, 2025), which was further analyzed through time-series methods. Resampling direction (PS→10 m, S2→3 m) minimally affects RMSE/MAE but significantly alters spatial structure and Moran's I values; downscaling PS to 10 m decreases Moran's I. M2 is suitable for standard applications, whereas M3 is preferable when preservation of spatial structure is important. Across the four examined scenarios, all methods reduce RMSE below the 0.07 agronomic threshold, with calibrated RMSE ranging from 0.02 to 0.05 (up to 0.06 across the full 56-pair dataset). M3's advantage lies in how effectively it reduces spatial autocorrelation mismatch: a 43.4% reduction in Moran's I (versus ~18.2% with M1 and M2) in the four example scenarios, and 39.5% versus 28.4% (M1) and 28.2% (M2) reduction over the full dataset.

1. Introduction

Rising costs and environmental concerns from excessive fertilizer use are encouraging the adoption of sustainable farming practices, including Variable Rate Application (VRA) technologies. VRA depends on prescription maps created from multispectral vegetation indices, especially the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Red Edge (NDRE), to optimize input application across varied fields (Ma et al., 2025). The Sentinel-2 (S2) constellation, operated by the European Space Agency under the Copernicus program, is widely considered the operational standard for agricultural remote sensing applications (Drusch et al., 2012). However, its spatial resolution limitations (especially the 20 m ground sampling distance of the Red Edge bands) and temporal constraints (a 5-day revisit period) pose challenges for Northern Italian rice farms characterized by small, irregularly shaped fields. The resulting mixed-pixel effect at field boundaries affects pixel purity and reduces the accuracy of within-field variability maps (Baldin and Casella, 2023). The PlanetScope (PS) constellation, operated by Planet Labs, addresses these spatial limitations with a 3 m ground sampling distance across all bands, including Red Edge (Frazier and Hemingway, 2021), and has demonstrated potential for within-field crop parameter estimation in combination with S2 (Raj et al., 2025). However, despite PlanetScope's radiometric harmonization of the visible and NIR bands, the PS Red Edge band (705 nm, FWHM ≈ 15 nm) remains unharmonized relative to S2 Band 5 (704.1 nm, FWHM ≈ 15 nm), resulting in systematic NDRE offsets that exceed the agronomic decision-making threshold. Analysis conducted between 2021 and 2023 at the study site shows consistent seasonal misalignment between PS and S2, with NDVI showing lower RMSE than NDRE (Baldin, 2025). The larger NDRE discrepancy is mainly due to differences in the Relative Spectral Response (RSR) functions of the PS Red Edge band and S2 Band 5, as well as different NIR band center wavelengths (PS: 865 nm vs S2 B8: 832.8 nm). Cross-sensor harmonization studies for Landsat-Sentinel pairs have demonstrated that both statistical and physically based calibration methods are feasible (Claverie et al., 2018; Roy et al., 2016). The agronomic miscalibration threshold is set at RMSE = 0.07 on the 0–1 normalized index scale, based on field-level sensitivity analysis of VRA prescription accuracy at the study site (Baldin and Casella, 2024) and should be validated independently at other sites. This threshold indicates the maximum NDRE error that results in minimal change to nitrogen prescription zones for rice cultivation. This study compares three intercalibration methods to determine the best approach for aligning PS NDRE with S2 reference values, focusing on both overall accuracy (RMSE, MAE) and spatial consistency (Moran's I). The comparison is conducted across two phenological stages (early and mid-season) and two resampling directions (PS→10 m and S2→3 m). This study also consolidates the methodological progression developed across the authors' previous work. Earlier investigations focused on index-domain intercalibration, first through a composite workflow combining linear regression, histogram shifting, and histogram matching, and then through a simplified histogram-matching-only alternative, when additional regression steps did not yield measurable advantages. More recent work extended the problem to spectral-band alignment prior to vegetation-index derivation. The present manuscript compares these three approaches within a single experimental framework, allowing both continuity with previous studies and a clearer assessment of the trade-off between global radiometric agreement and preservation of spatial structure.

2. Methodology

2.1 Data and Preprocessing

S2 Level-2A bottom-of-atmosphere (BOA) surface reflectance products were obtained from the Copernicus Data Space Eco-system. The selected bands include Red (B4, 664.6 nm, 10 m), Red Edge 1 (B5, 704.1 nm, 20 m), and NIR (B8, 832.8 nm, 10

https://doi.org/10.5194/isprs-archives-XLIX-B1-2026-231-2026 |

m). The Level-2A atmospheric correction is performed using the Sen2Cor processor, which performs radiometric calibration, atmospheric correction using lookup tables, and terrain correction. PlanetScope SuperDove (PSB.SD) surface reflectance products were acquired from Planet Labs through an academic license. The bands used are: Red (B6, 665 nm, 3 m), Red Edge (B7, 705 nm, 3 m), and NIR (B8, 865 nm, 3 m). Planet’s SR product includes atmospheric correction and radiometric harmonization with S2 for visible and NIR bands; however, the Red Edge band remains unharmonized, which is the main source of NDRE discrepancy. Both products are delivered in UTM/WGS84 projection with 12-bit radiometric resolution stored as 16-bit unsigned integers. Cloud-free scenes were selected using visual inspection and quality-assessment metadata. Temporal co-registration was achieved by choosing pairs of images acquired on the same day or within a close timeframe, based on the acquisition-screening criteria outlined in Baldin (2025). This process resulted in 56 pairs of S2 and PS images.

2.2 Spatial Co-Registration and AOI Masking

Since PS and S2 have different ground sampling distances, spatial co-registration is necessary before pixel-by-pixel comparison. Two resampling methods are tested: upscaling S2 (original spatial resolutions: 10 m for Red, 20 m for Red Edge) to the PS 3 m grid, and downscaling PS (all original bands at 3 m) to the S2 10 m resolution grid. Nearest-neighbor interpolation is used for all main analyses and results reported in this paper. Additional resampling methods (bilinear, cubic, makima) were tested in a sensitivity analysis, resulting in negligible differences relative to the defined threshold (RMSE = 0.07) and are therefore not discussed further. The Area of Interest (AOI) mask is created from a field boundary Shapefile with a 10 m inward negative buffer (polybuffer with inward erosion) to remove mixed-pixel effects at field edges. All normalization and statistical calculations are performed strictly within this masked AOI, using a three-way intersection: target pixels are not NaN, reference pixels are not NaN, and pixels are within the buffered AOI. Valid pixels are extracted through a three-way intersection mask defined as follows:

- Target pixels are not NaN
- Reference pixels are not NaN
- Pixels are within the Area of Interest (AOI)

$$\mathrm{Mask}(i,j) = \mathrm{AOI}(i,j) \wedge \neg\,\mathrm{NaN}_{\mathrm{target}}(i,j) \wedge \neg\,\mathrm{NaN}_{\mathrm{reference}}(i,j) \quad (1)$$

2.3 Method Numbering

In this manuscript, the methods are defined as follows: M1 combines linear regression, histogram shifting and histogram matching; M2 uses only histogram matching; M3 performs band-level alignment prior to index computation in accordance with the methodology previously presented (Baldin and Casella, 2025a).

2.4 Method 1: Linear Regression + Histogram Shifting + Histogram Matching (M1)

M1 is a three-stage composite pipeline that integrates parametric modeling with nonparametric refinement (Baldin and Casella, 2024). Stage 1 applies a global ordinary least squares (OLS) regression to valid pixel pairs within the AOI, generating gain and offset coefficients that correct the primary linear component of the sensor discrepancy. Stage 2 calculates the difference between the regression output and the reference, then shifts the corrected distribution to recenter it. Stage 3 uses histogram matching to adjust residual non-linear distributional differences. The rationale for M1 is to decompose the calibration problem into gain/offset correction, mean recentering, and distribution-shape refinement, which may improve stability when distributions have significantly different means or spreads (when multiple nonlinear discrepancies coexist between the target and reference distributions).

2.4.1 Global Linear Regression

An ordinary least squares (OLS) model is fitted to valid pixel pairs within the AOI, which results in the regression-predicted reference value:

$$\hat{R}(i,j) = a \cdot T(i,j) + b \quad (2)$$

Where the other terms are:

- “T(i,j)” is the target pixel value,
- “a” is the gain (adjusts sensor spectral response and atmospheric transmittance),
- “b” is the offset (accounts for additive differences like atmospheric path radiance).

This model is applied globally to produce the first intermediate output.

2.4.2 Histogram Average Shifting

Compute the mean-shift:

$$\Delta\mu = \mathrm{mean}\Big(\hat{R}_{\mathrm{valid}} - T_{\mathrm{valid}}\Big) \quad (3)$$

Subtract Δμ from the regression prediction to recenter the corrected distribution:

$$R_{\mathrm{shifted}}(i,j) = \hat{R}(i,j) - \Delta\mu \quad (4)$$

2.4.3 Final Histogram Matching

The shifted image is further refined using Histogram Matching (HM), as described step-by-step in the next paragraph for M2, to correct residual non-linear distributional differences:

$$R_{\mathrm{final}}(i,j) = \mathrm{HM}(R_{\mathrm{shifted}}(i,j)) \quad (5)$$

where the second part of the equation denotes the histogram matching operation, mapping the cumulative distribution function (CDF) of the shifted image to the reference CDF.

2.5 Method 2: Histogram Matching (M2)

M2 uses histogram matching (HM), a non-parametric radiometric normalization method that adjusts the cumulative distribution function (CDF) of the PlanetScope NDRE to match the S2 reference CDF (Gonzalez and Woods, 2018; Schroeder et al., 2006). Valid pixel values are normalized to the [0,1] range using shared global extrema. The CDF transfer function maps each quantile of the target distribution to the corresponding quantile of the reference distribution, implemented through MATLAB’s imhistmatch() function. The matched values are then denormalized back to the original radiometric domain, and the image is reconstructed, preserving NaN values outside the AOI. M2 is computationally efficient, requires no parametric assumptions, and operates directly on the vegetation index without requiring access to individual spectral bands (Baldin and Casella, 2025b). Histogram matching is implemented through the following steps:

2.5.1 Global Normalization

Normalize valid pixel values to the range between 0 and 1 using shared global extrema:

https://doi.org/10.5194/isprs-archives-XLIX-B1-2026-231-2026 | © Author(s) 2026. CC BY 4.0 License.

$$x_{\text{norm}}(i,j) = \frac{x(i,j) - x_{\min}}{x_{\max} - x_{\min}} \tag{6}$$

where x(i,j) is the pixel value, x_{min} and x_{max} are the minimum and maximum values across both images.

2.5.2 CDF Matching

Align the CDF of the target image with the reference using a transfer function.

$$F_{\text{target}}(x) \longrightarrow F_{\text{reference}}(x) \tag{7}$$

The transfer function T maps each quantile in the target to the corresponding quantile in the reference:

$$x' = T(x) \tag{8}$$

This is implemented through MATLAB's imhistmatch() with default parameters.

2.5.3 Denormalization

Convert normalized values back to the original radiometric domain:

$$x_{\text{denorm}}(i,j) = x'_{\text{norm}}(i,j) \cdot (x_{\max} - x_{\min}) + x_{\min} \tag{9}$$

2.5.4 Image Reconstruction

Reconstruct the image, preserving NaN values outside the AOI:

$$x_{\text{final}}(i,j) = \begin{cases} x_{\text{denorm}}(i,j), & \text{if Mask}(i,j) = 1 \\ \text{NaN}, & \text{otherwise} \end{cases} \tag{10}$$

2.6 Method 3: Band Alignment (M3)

M3 addresses the non-uniform scaling issue between sensors by performing calibration at the spectral band level before computing the final index (Baldin and Casella, 2025a). Independent OLS linear regression is applied to each band (Red Edge, NIR) using valid PS and S2 pixel pairs within the AOI to determine gain and offset. Calibrated NDRE is then recalculated from the adjusted bands. Normalizing each band before index calculation maintains the physical inter-band relationships set by the reference sensor, which is essential for vegetation indices given the biophysical significance of the NIR-to-Red Edge reflectance ratio.

2.6.1 Per-band OLS Linear Regression

For each band (Red Edge, NIR) the regression- predicted reference value for band b is:

$$\hat{S}_b(i,j) = a_b \cdot P_b(i,j) + b_b \tag{11}$$

Where the other terms are:

- $P_b(i,j)$ is the target pixel value for band b,
- a_b is the gain for band b,
- b_b is the offset for band b.

2.6.2 Index Calculation from Aligned Bands

After aligning all bands, the spectral index (NDRE) is recalculated using the raster calculator formula:

$$\text{Index}(i,j) = \frac{\hat{S}_{B_1}(i,j) - \hat{S}_{B_2}(i,j)}{\hat{S}_{B_1}(i,j) + \hat{S}_{B_2}(i,j)} \tag{12}$$

where the aligned bands are used in the formula.

2.6.3 Single-band Features

For single-band features, M3 reduces to a single OLS model applied directly:

$$\hat{S}(i,j) = a \cdot P(i,j) + b \tag{13}$$

2.7 Statistical Evaluation Framework

After applying each normalization method, a set of metrics is calculated within the buffered AOI. Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) measure the overall difference between calibrated PS and reference S2 NDRE values. Moran's I (Moran, 1950) is calculated using a Queen's contiguity weight matrix to evaluate spatial autocorrelation of the PS image, the reference S2 image, and especially the difference map (PS − S2). A high Moran's I on the difference map indicates clustered residual errors, while values near zero suggest random residuals, signaling effective calibration. The implementation employs 2D convolution with the Queen's kernel for efficiency, excluding NaN pixels from both the numerator and denominator. The percentage decrease in Moran's I for the difference map (PS − S2), compared to the uncalibrated case, shows how well each method reduces residual spatial autocorrelation while keeping the reference image's spatial structure. For multiple acquisition dates, overall reduction is the arithmetic mean of daily percentage reductions.

2.7.1 Global Comparative Metrics

Root Mean Square Error (RMSE):

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(T_i - R_i)^2} \tag{14}$$

Where T_i and R_i are the target and reference pixel values, and N is the number of valid pixels.

Mean Absolute Error (MAE):

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|T_i - R_i| \tag{15}$$

2.7.2 Spatial Autocorrelation: Moran's I

$$I = \frac{N}{W} \cdot \frac{\sum_i \sum_j w_{ij}\,(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2} \tag{16}$$

where:

- N = number of valid pixels,
- x_i = value at pixel i,
- $\bar{x}$ = mean of all x_i,
- w_{ij} = spatial weight between pixels i and j (using a 3x3 Queen's contiguity weight matrix)

$$W = \sum_i \sum_j w_{ij} \tag{17}$$

Moran's I is applied to the reference image, the target image, and to the difference map (PS − S2) to assess whether residual errors exhibit spatial clustering (I → +1) or approximate spatial randomness (I ≈ 0). The implementation uses 2D convolution with the Queen’s kernel for computational efficiency, with NaN pixels contributing neither to the numerator nor the denominator.

3. Study Area

The test site is “Riserva San Massimo,” a rice farm located in Gropello Cairoli (Pavia), within the Ticino Valley Park in Northern Italy (Figure 1). The reserve covers about 800 hectares, primarily dedicated to rice farming. Precision techniques are used, with nitrogen applied variably based on satellite and drone vegetation indices. Five rice varieties are grown in 83 fields, many of which are small and irregular, creating challenges for edge effects in 10–20 m-resolution imagery. The site has served as a test area for PS-S2 interoperability research since 2021 (Baldin and Casella, 2023, 2024; Baldin, 2025). The ArcGIS Pro layout (Figure 1) is in EPSG:4326, with top-left pixel at 8° 57′ 46″ E, 45° 12′ 39″ N.

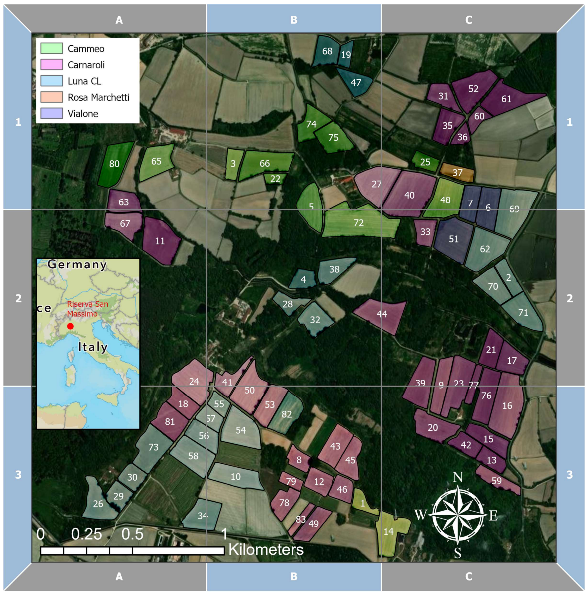

Figure 1. PS image (11 May 2022) of “Riserva San Massimo”

4. Results

Tables 1–4 present four NDRE intercalibration scenarios from 56 image pairs (2021–2023), covering two resampling directions (PS to 10 m, S2 to 3 m) and two phenological stages (early-season: 11 May 2022; mid-season: 15 July 2022). May reflects early canopy development, with sparse vegetation and a significant water/soil background; July represents peak canopy closure, with a more uniform spectral response. Uncalibrated PS NDRE consistently exceeded the 0.07 RMSE agronomic threshold across all four scenarios, with RMSE values ranging from 0.091 (July) to 0.188 (May). The RMSE and MAE values are nearly equal in the May scenario (0.188 vs 0.180), indicating that the error is predominantly systematic (bias) rather than random, confirming that the miscalibration is correctable through distribution alignment. In all four tested scenarios, all three methods reduced RMSE below the 0.07 agronomic threshold. M1 and M2 produced similar RMSE and MAE values across these four scenarios, although small differences remain visible in the full 56-date time series. M3 achieved the strongest Moran’s I gap reduction across all scenarios (average 43.4% vs 18.2% for M1/M2), indicating a substantially better correction of spatial autocorrelation differences between sensors. In the July scenarios, M3 also achieved the lowest RMSE, demonstrating that band-level alignment yields both improved spatial fidelity and higher global accuracy under full-canopy conditions.

PS → 10 m	**Moran’s I**		**PS vs. S2**	
PS Calibration	**PS**	**PS - S2**	**RMSE**	**MAE**
None	0.8057	0.4343	0.1880	0.1804
M1	0.7371	0.2904	0.0428	0.0307
M2	0.7374	0.2905	0.0427	0.0307
M3	0.6158	0.0875	0.0448	0.0343

Table 1. NDRE, 11/05/2022; S2 Moran’s I: 0.7199

S2 → 3 m	**Moran’s I**		**PS vs. S2**	
PS Calibration	**PS**	**PS - S2**	**RMSE**	**MAE**
None	0.8926	0.7260	0.1871	0.1795
M1	0.8527	0.6576	0.0441	0.0316
M2	0.8530	0.6580	0.0441	0.0316
M3	0.7798	0.5236	0.0454	0.0347

Table 2. NDRE, 11/05/2022; S2 Moran’s I: 0.8708

PS→10 m	**Moran’s I**		**PS vs. S2**	
PS Calibration	**PS**	**PS - S2**	**RMSE**	**MAE**
None	0.9283	0.5687	0.0910	0.0869
M1	0.9237	0.4474	0.0261	0.0202
M2	0.9235	0.4474	0.0261	0.0202
M3	0.9205	0.2999	0.0219	0.0165

Table 3. NDRE, 15/07/2022; S2 Moran’s I: 0.9218

S2 → 3 m	**Moran’s I**		**PS vs. S2**	
PS Calibration	**PS**	**PS - S2**	**RMSE**	**MAE**
None	0.9747	0.7601	0.0915	0.0872
M1	0.9724	0.6922	0.0271	0.0208
M2	0.9724	0.6918	0.0271	0.0208
M3	0.9729	0.6180	0.0231	0.0173

Table 4. NDRE, 15/07/2022; S2 Moran’s I: 0.9705

Uncalibrated errors are greater early in the season (May) than mid-season (July), reflecting increased spectral variability caused by water and soil backgrounds under sparse canopies. Resampling direction affects spatial autocorrelation: upscaling S2 to 3 m replicates each pixel across many smaller cells, artificially raising Moran’s I, while downscaling PS to 10 m aggregates variable pixels, lowering it. As established in the literature, Moran’s I increases when nominal resolution is increased without adding real information (Moran, 1950). Table 5 reports normality diagnostics for the calibrated residual distributions for the Table 1 scenario (11/05/2022, PS→10 m).

Statistic	**M1**	**M2**	**M3**
Mean (μ)	-0.0004	0.0000	0.0000
Std (σ)	0.0628	0.0702	0.0474
Median	0.0036	-0.0065	-0.0048
Skewness	-1.0689	0.5026	0.9667
Kurtosis	7.7400	6.6141	7.4926
K–S p-value	0.000000	0.000000	0.000000

Table 5. Normality diagnostics for calibrated residuals (11/05/2022, PS resampled to 10 m, N = 18,587)

After calibration, each method yields residuals close to zero, indicating successful bias removal. However, dispersion differs across methods: M3 achieves the lowest standard deviation (σ = 0.0474), followed by M1 (σ = 0.0628) and M2 (σ = 0.0702). Despite the near-zero means, all residual distributions remain non-normal (K–S test, $p < 0.0001$), with persistent skewness and leptokurtic behavior (kurtosis 6.6–7.7) consistent with mixed-pixel and field-boundary effects in the underlying remote-sensing data. The residual statistics in Table 5 are computed in each method's internal processing domain and are therefore not directly comparable to the RMSE and MAE values in Tables 1–4, which are computed from pixelwise NDRE differences in the original radiometric domain. A graphical comparison of the NDRE analysis pipeline for the worst-case scenario (May 2022, PS resampled to 10 m to reduce spatial autocorrelation) is shown in Figures 2–9. The workflow includes NDRE maps, raster differences, histograms, and scatter plots. For the map panels (EPSG:32632), the left panel shows the S2 reference NDRE, the right panel shows the PS original or aligned NDRE, and the bottom-left panel shows the difference raster between them. Each graphical panel is accompanied by the corresponding RMSE, MAE, and Moran’s I values to support visual and quantitative comparison. Figures 10-12 display residual histograms and Q–Q plots, which corroborate the findings detailed in Table 5. All aligned distributions remain non-normal, as indicated by the Kolmogorov–Smirnov test and Q–Q plots, with persistent skewness and leptokurtic behavior. A complementary time-series analysis for all 56 image pairs is shown in Figures 13-18, using the same pipeline with PS downscaled to 10 m by nearest-neighbor interpolation. Table 6 summarizes the 56-date time-series results. The mean reduction percentages reported reflect the average of the daily performance metrics. All methods remained below the 0.07 threshold (56/56 dates), with M3 achieving the lowest RMSE in 43/56 dates and the lowest difference-map Moran’s I in 33/56 dates. M3 also exhibited the lowest maximum RMSE (0.0526 vs 0.0601/0.0603). M3’s stronger reduction in difference-map Moran’s I is visible in Figures 6 and 8, where the residual field after M3 calibration is less structured than after M2 calibration, especially in field 82.

56-date statistics	M1	M2	M3
RMSE (post-calibration)			
Mean	0.0337	0.0338	0.0317
Std	0.0107	0.0108	0.0097
Max	0.0601	0.0603	0.0526
Dates < 0.07	**56/56**	**56/56**	**56/56**
MAE (post-calibration)			
Mean	0.0244	0.0244	0.0233
Max	0.0459	0.0470	0.0389
Moran’s I (PS−S2 difference map)			
Mean (pre)	0.3579	0.3579	0.3579
Mean (post)	0.1939	0.1942	0.1776
Mean of daily reductions	28.4%	28.2%	39.5%
Method comparison (dates out of 56)			
Lowest RMSE	10/56	3/56	**43/56**
Lowest Moran’s I (difference map)	16/56	7/56	**33/56**

Table 6. Summary statistics across the 56-date dataset (PlanetScope resampled to 10 m, 2021–2023)

The “Mean of daily reductions” averages percentage drops across 56 dates, not the total shift from Mean (pre) → (post).

Figures 2–18 were generated using MATLAB R2025b and the authors’ toolbox (Baldin and Casella, 2026).

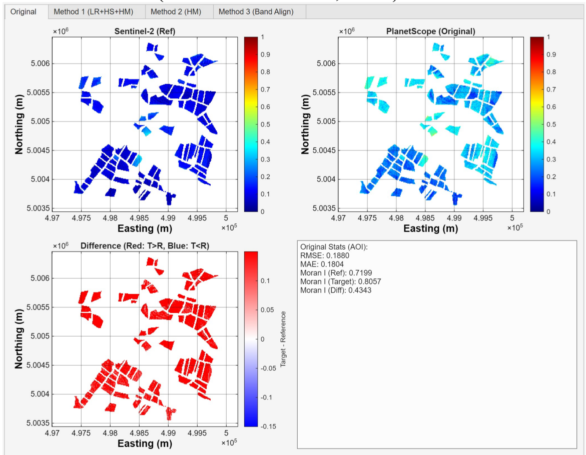

Figure 2. NDRE Maps comparison without calibration

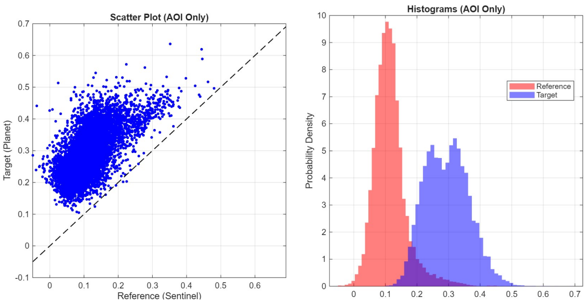

Figure 3. NDRE Scatter Plot and Histogram without calibration

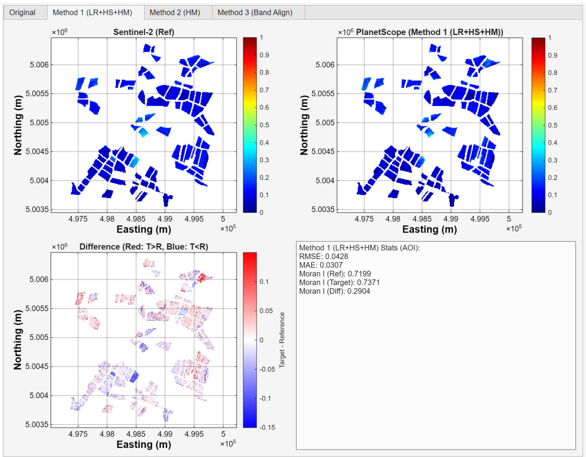

Figure 4. NDRE Maps after M1 alignment (LR+HS+HM)

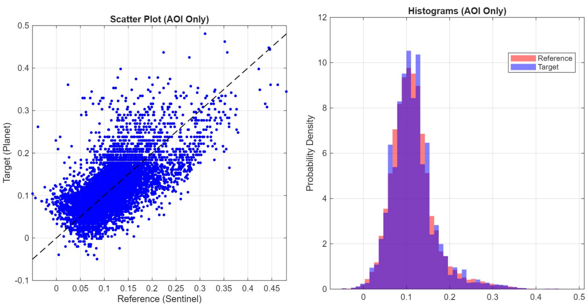

Figure 5. NDRE Scatter Plot and Histogram after M1

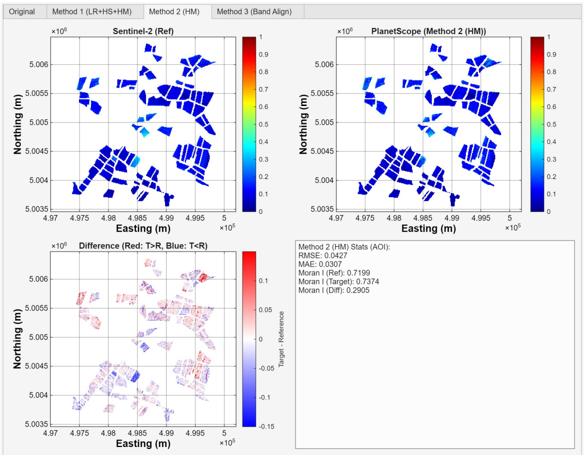

Figure 6. NDRE Maps after M2 alignment

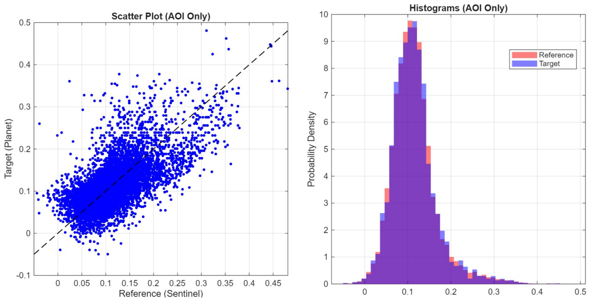

Figure 7. NDRE Scatter Plot and Histogram after M2

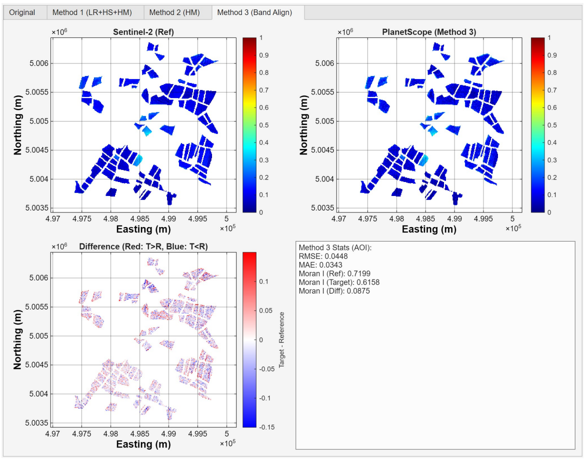

Figure 8. NDRE Maps after M3 alignment

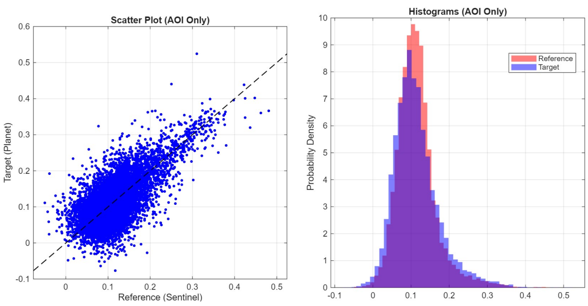

Figure 9. NDRE Scatter Plot and Histogram after M3

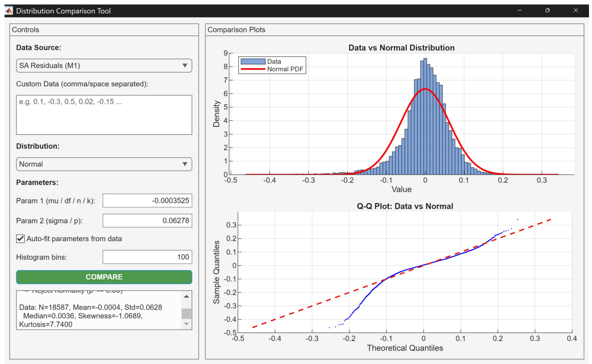

Figure 10. Residual histograms and Q–Q plots for M1

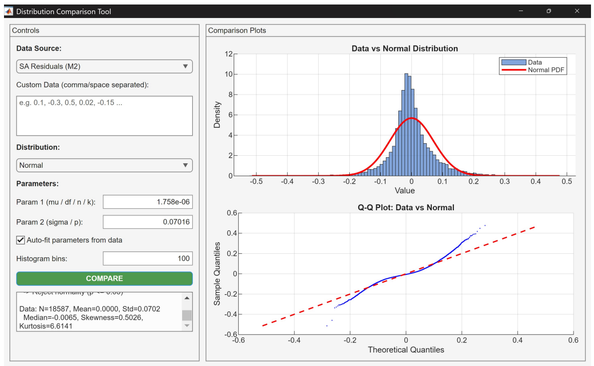

Figure 11. Residual histograms and Q–Q plots for M2

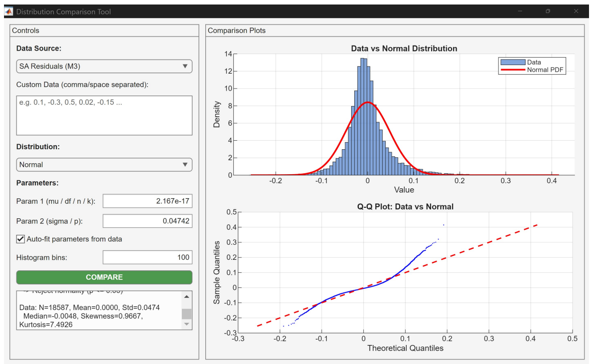

Figure 12. Residual histograms and Q–Q plots for M3

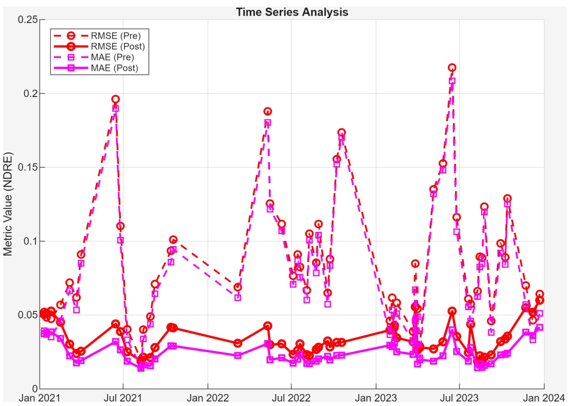

Figure 13. RMSE and MAE across 56 dates for M1

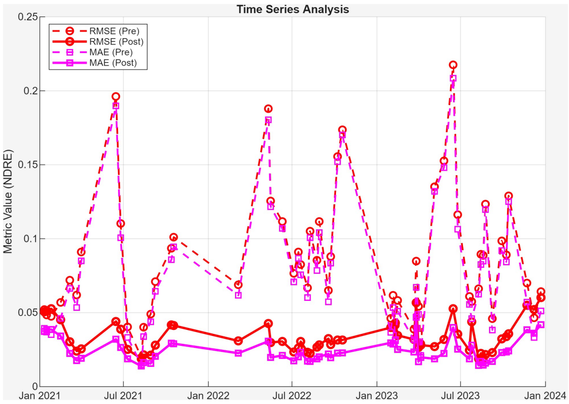

Figure 14. RMSE and MAE across 56 dates for M2

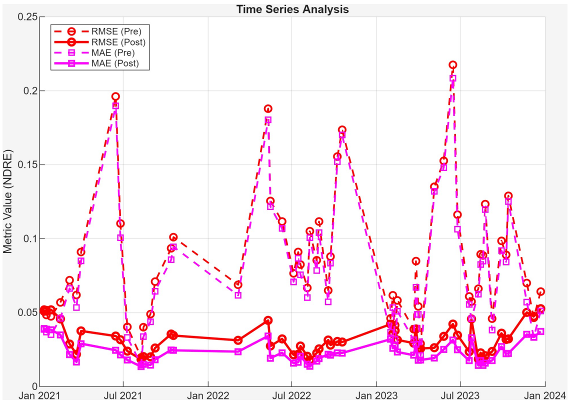

Figure 15. RMSE and MAE across 56 dates for M3

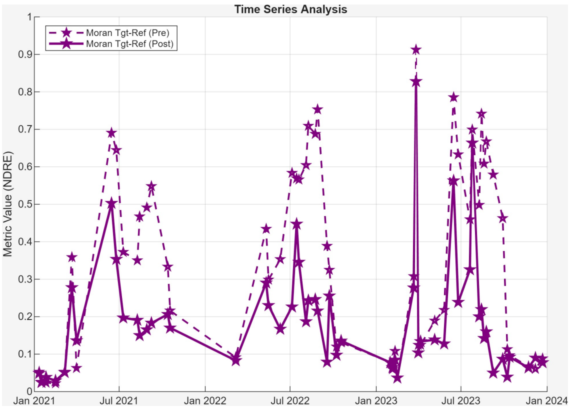

Figure 16. Moran’s I across 56 dates for M1

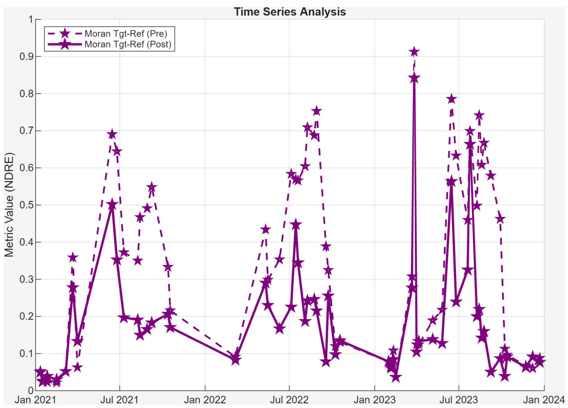

Figure 17. Moran’s I across 56 dates for M2

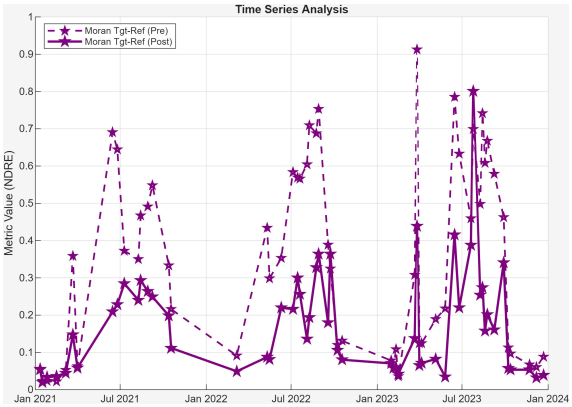

Figure 18. Moran’s I across 56 dates for M3

5. Discussion

All three calibration methods effectively eliminated the global NDRE offset across the four analyzed scenarios and the 56-date time series, reducing post-calibration RMSE to 0.02–0.05 (up to 0.06 over the full dataset), well below the 0.07 agronomic threshold. This RMSE is low relative to the agronomic decision threshold adopted in this study and indicates strong cross-sensor agreement after calibration. Calibration and evaluation are performed on the same scene, consistent with standard radiometric-normalization practice (Claverie et al., 2018; Schroeder et al., 2006), while the 56-date time series provides temporal generalization across independent acquisition dates. These results indicate that calibrated PS can support VRA prescription map generation in a manner comparable to S2 at the study site, without significantly affecting nitrogen management decisions under the tested conditions. The methods differ primarily in how they handle spatial structure. M1 and M2, which operate in the index domain, force the NDRE distribution to match the reference, correcting global bias and variance but not band-level discrepancies: S2 yields higher reflectance than PS in the Red and Red Edge bands, while PS exceeds S2 in NIR reflectance. By correcting each band independently before index computation, M3 preserves the physical inter-band ratio structure that carries biophysical meaning in normalized difference indices. This mechanism explains M3’s greater reduction in Moran’s I (39.5% vs 28.4%/28.2% for M1/M2): band-level alignment removes spatially structured errors that index-level matching cannot address. The preserved Moran’s I of the calibrated PS image (Tables 1–4, column 2), combined with the reduced Moran’s I of the difference map, confirms that spatial structure is largely maintained in the signal while being attenuated in the residual error. Improvements in spatial autocorrelation and RMSE are complementary but independent, since RMSE averages error magnitude while Moran’s I captures spatial arrangement. M1 did not deliver operational advantages over M2, confirming that the additional regression and shifting stages are unnecessary when target and reference distributions share similar shapes. Testing M1 remains warranted for datasets with stronger offsets or non-contemporaneous acquisitions (Baldin and Casella, 2024). The seasonal asymmetry in uncalibrated error (higher in May than in July) reflects the greater influence of soil and water backgrounds under sparse vegetation, compressing the NDRE dynamic range and amplifying cross-sensor RSR differences. The rejection of normality in all calibrated residual distributions (Table 5) should not be interpreted as a calibration failure. Persistent leptokurtic behavior is expected in remote-sensing data affected by mixed pixels and field boundaries, and distributional diagnostics should complement RMSE, MAE, and Moran's I as evaluation criteria. All three methods assume a spatially stationary PS-to-S2 relationship across the AOI. Spatially adaptive calibration strategies could be explored in future work. This paper specifically assesses NDRE as the principal index for nitrogen prescription; however, the employed toolbox (Baldin and Casella, 2026) applies the same calibration methodology to additional vegetation indices, including NDVI, SAVI, EVI, and user-defined indices. Methods M1 and M2 function within the index domain, while M3 performs calibration at the band level via per-band ordinary least squares regression. The toolbox provides capabilities for calibration, histogram matching, shifting, and statistical summarization in both single-date and time-series contexts, along with support for user-defined formulas and per-band alignment. The main limitation remains single-site validation at "Riserva San Massimo", which could be addressed through multi-site testing.

6. Conclusions

Each of the three methods effectively reduced the initial miscalibration to below the agronomic threshold of 0.07 RMSE across all 56 dates spanning three growing seasons (2021–2023), with mean post-calibration RMSE of 0.034 (M1), 0.034 (M2), and 0.032 (M3). These values confirm that PS NDRE can be reliably aligned with S2. M1 and M2 produced nearly identical accuracy, confirming that histogram matching alone is adequate when computational simplicity is prioritized. M3 achieved the lowest RMSE in 43 of 56 dates and the lowest difference-map Moran's I in 33 of 56 dates, with the best worst-case performance (maximum RMSE 0.0526 vs 0.0601/0.0603 for M1/M2). Residual distributions remain non-normal after all calibrations, indicating that distributional diagnostics should complement RMSE and Moran's I as evaluation criteria. The findings support a practical decision framework: histogram-based alignment (M2) for rapid operational interoperability, and band-level alignment prior to index calculation (M3) for applications requiring spatial fidelity, such as edge-effect correction (Baldin, 2025). Single-site validation is the main limitation and can be resolved with toolbox-supported multi-site testing (Baldin and Casella, 2026).

Acknowledgements

The authors thank Dr. Agr. G. L. Rognoni and Ing. C. Bergonzi for their expertise in site selection and field data interpretation. Appreciation is also extended to the owners and staff of "Riserva San Massimo" for sharing their experience and support. This contribution is part of the NODES project, funded by the European Union - NextGenerationEU, Mission 4 Component 1.5 - ECS00000036 - CUP F17G22000190007.

References

Baldin, C.M., 2025. The use of high resolution and high frequency satellite images as a support for precise and sustainable agriculture. PhD Thesis, University of Pavia. https://hdl.handle.net/11571/1530777

Baldin, C.M., Casella, V.M., 2023. Comparison between the vegetation indices obtained from Sentinel-2 and Planet: a case study over a rice farm in Northern Italy. In: Italian Conference on Geomatics and Geospatial Technologies, Springer, pp. 257–284. https://doi.org/10.1007/978-3-031-59925-5_18

Baldin, C.M., Casella, V.M., 2024. Vegetative index intercalibration between PlanetScope and Sentinel-2 through a SkySat classification in the context of "Riserva San Massimo" rice farm in Northern Italy. Remote Sens., 16(21), 3921. https://doi.org/10.3390/rs16213921

Baldin, C.M., Casella, V.M., 2025a. Comparison of PlanetScope and Sentinel-2 spectral channels and their alignment via linear regression for enhanced index derivation. Geosciences, 15(5), 184. https://doi.org/10.3390/geosciences15050184

Baldin, C.M., Casella, V.M., 2025b. Assessment of PlanetScope imagery for the "Riserva San Massimo" rice farm: evaluating benefits and calibration in relation to Sentinel-2 data. In: Italian Conference on Geomatics and Geospatial Technologies, Springer, pp. 179–197. https://doi.org/10.1007/978-3-031-91144-6_12

Baldin, C.M., Casella, V.M., 2026. Remote Sensing Toolbox (Version 3). Zenodo. https://doi.org/10.5281/zenodo.19207153

Claverie, M., Ju, J., Masek, J.G., Dungan, J.L., Vermote, E.F., Roger, J.-C., Skakun, S.V., Justice, C., 2018. The Harmonized Landsat and Sentinel-2 surface reflectance data set. Remote Sens. Environ., 219, 145–161. https://doi.org/10.1016/j.rse.2018.09.002

Drusch, M., Del Bello, U., Carlier, S., Colin, O., Fernandez, V., Gascon, F., Hoersch, B., Isola, C., Laberinti, P., Martimort, P., Meygret, A., 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. Remote Sens. Environ., 120, 25–36. https://doi.org/10.1016/j.rse.2011.11.026

Frazier, A.E., Hemingway, B.L., 2021. A technical review of Planet smallsat data: practical considerations for processing and using PlanetScope imagery. Remote Sens., 13(16), 3087. https://doi.org/10.3390/rs13163087

Gonzalez, R.C., Woods, R.E., 2018. Digital Image Processing, 4th ed. Pearson, New York.

Ma, Y., Gong, J., Jin, T., Xu, T., Kan, G., 2025. Comparison of different vegetation indices for estimating vegetation changes and analyzing driving factors in a semi-arid area, China. J. Arid Land, 17, 1785. https://doi.org/10.1007/s40333-025-0083-x

Moran, P.A.P., 1950. Notes on continuous stochastic phenomena. Biometrika, 37(1/2), 17–23. https://doi.org/10.2307/2332142

Raj, R., Bayat, B., Raza, A., Gaiser, T., Vereecken, H., Montzka, C., 2025. Within-field crop leaf area index simulation using a hybrid PROSAIL-SVR approach: evaluating Sentinel-2 and PlanetScope potential. Int. J. Remote Sens., 46(22), 8546–8566. https://doi.org/10.1080/01431161.2025.2571231

Roy, D.P., Kovalskyy, V., Zhang, H.K., Vermote, E.F., Yan, L., Kumar, S.S., Egorov, A., 2016. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. Remote Sens. Environ., 185, 57–70. https://doi.org/10.1016/j.rse.2015.12.024

Schroeder, T.A., Cohen, W.B., Song, C., Canty, M.J., Yang, Z., 2006. Radiometric correction of multi-temporal Landsat data for characterization of early successional forest patterns in western Oregon. Remote Sens. Environ., 103(1), 16–26. https://doi.org/10.1016/j.rse.2006.03.008