Open Access

Earth System Science Data Discussions

Improving the usability of the MISR L1B2 Georectified Radiance Product (2000–present) in land surface applications

Michel M. Verstraete[1], Linda A. Hunt[2], and Veljko M. Jovanovic[3]

[1]Global Change Institute (GCI), University of the Witwatersrand, Braamfontein, Republic of South Africa.
[2]Science Systems and Applications, Inc. (SSAI), Hampton, VA 23666-5845, USA.
[3]NASA Jet Propulsion Laboratory (JPL), Pasadena, CA 91109, USA.

Correspondence: Michel M. Verstraete (Michel.Verstraete@wits.ac.za or MMVerstraete@gmail.com)

Abstract. The Multi-angle Imaging SpectroRadiometer (MISR) instrument on NASA's Terra platform has been acquiring global measurements of the spectro-directional reflectance of the Earth since 24 February 2000 and is still operational as of this writing. The primary radiometric data product generated by this instrument is known as the Level 1B2 Georectified Radiance Product (GRP): it contains the 36 radiometric measurements acquired by the instrument's 9 cameras, each observing the planet in 4 spectral bands. The product version described here is projected on a digital elevation model and is available from the NASA Langley Atmospheric Science Data Center (ASDC) (`https://doi.org/10.5067/Terra/MISR/MI1B2T_L1.003`; Jovanovic et al. (1999)). The MISR instrument is highly reliable. Nevertheless, its on-board computer occasionally becomes overwhelmed by the amount of raw observations coming from the cameras' focal planes, especially when switching into or out of Local Mode acquisitions that are often requested in conjunction with field campaigns. Whenever this occurs, one or more lines of data are dropped while the computer resets and readies itself for accepting new data. Although this type of data loss is minuscule compared to the total amount of measurements acquired, and is marginal for atmospheric studies dealing with large areas and long periods of time, this outcome can be crippling for land surface studies that focus on the detailed analysis of particular scenes at specific times. This paper describes the problem, reports on the prevalence of missing data, proposes a practical solution to optimally estimate the values of the missing data and provides evidence of the performance of the algorithm through specific examples in South Africa. The software to process MISR L1B2 GRP data products as described here is openly available to the community from the GitHub web site `https://github.com/mmverstraete` or `https://doi.org/10.5281/zenodo.3519989`.

1 Introduction

The Multi-angle Imaging SpectroRadiometer (MISR) is one of the five Earth observing instruments hosted on the NASA Terra platform, launched into a sun-synchronous low-Earth orbit on 18 December 1999 (altitude: 713 km; equatorial crossing time: 10:30 AM, local time). It features 9 cameras pointing at various bore angles, both forward and aft of the platform,

and each camera collects data in 4 spectral bands (blue, green, red and near-infrared), for a total of 36 data channels. MISR became operational on 24 February 2000 and is still collecting data as of this writing. This instrument has been extensively
25   described in the literature (Diner et al. (1989); Diner et al. (1998); Diner et al. (1999a); Diner et al. (2002); Diner et al. (2005)), and more recently in a special issue of the MDPI journal *Remote Sensing* dedicated to that instrument (See https://www.mdpi.com/journal/remotesensing/special_issues/misr_rs).

The primary mission of the MISR instrument is to quantitatively characterize the state and evolution of the Earth's atmosphere (clouds and aerosols) and surface (vegetation and soils over land, or water properties over oceans). Specifically,
30   its combined multi-angular and multi-spectral capability has proven very useful to investigate land surface processes, especially when coupled with in situ observations (field campaigns) or other sources of environmental data (e.g., Liu et al. (2018); Mahlangu et al. (2018); Van den Hoof et al. (2018)). In its default Global Mode of operation, the data downloaded from the satellite to the ground segment are provided either at the native spatial resolution of the instrument (275 m) or at a reduced spatial resolution (1100 m), depending on the data channel, to reduce the overall data rate. However, the instrument can occa-
35   sionally be operated in Local Mode, whereby all 36 data channels are made available at the finer spatial resolution of 275 m, albeit for a limited region, typically about 300 km along-track.

It turns out that when the MISR instrument is switched from Global to Local Mode (and conversely), the on-board computer may sometimes become overwhelmed by the amount of data coming from the focal planes of the various cameras. When this happens, the computer may drop one or more lines of measurements while it resets itself and proceeds with the processing
40   of subsequent lines. Such events can also occasionally occur outside of this Mode switching context; they often perturb most spectral bands of a camera, and more than one camera. These missing lines are detected as part of the Level 1 processing; they are marked with the appropriate code indicating that measurements were lost and all data products subsequently derived from these sources are of course unavailable too.

These events imply significant consequences for land surface products, which often rely on multiple spectral bands, e.g.,
45   vegetation indices, or on the whole range of cameras and bands, e.g., in assessing advanced products like surface albedo and the Fraction of Absorbed Photosynthetically Active Radiation (FAPAR), for instance. Such data losses can be crippling to the investigation, especially if they occur systematically in space and time, or at the very place and time when a field campaign is taking place.

A procedure is thus required to minimize the number of missing values in MISR data files, and to maximize the spatial
50   coverage of derived land surface products. The purpose of this paper is to report on the prevalence of this problem, to describe the proposed solution, and to show the performance of the suggested approach through concrete examples.

2   Problem statement and motivation

The primary MISR data, from which all other atmospheric and surface products are derived, are the L1B2 Georectified Radiance Products (GRP), which contain, in particular, the terrain-projected product of interest here. The latter provides
55   calibrated multi-angular and multi-spectral observations of the planetary reflectance measured by the instrument at the nominal

'Top of the Atmosphere' (ToA), projected to a Space-Oblique Mercator (SOM) grid and geo-referenced on a digital elevation model.

2.1 Global and Local Modes of operation

L1B2 GRP radiance measurements in each of the 9 cameras and 4 spectral bands are acquired by the instrument at the native
60   spatial resolution of 275 m (across-track, for the 8 off-nadir pointing cameras), or 250 m (across-track, for the nadir-pointing camera), although the latter is subsequently resampled to 275 m to provide the same sampling in all full resolution data channels. The resulting instrument-wide rate of data acquisition exceeds the downlink bandpass allotted to MISR by the Terra platform. To fit that constraint in its default Global Mode of operation, 12 of those 36 data channels (the 4 spectral bands of the nadir-pointing camera and the red spectral band of the 8 off-nadir pointing cameras) are transmitted to the ground segment at
65   the full, native spatial resolution, while the remaining 24 data channels (i.e., the non-red spectral bands of the off-nadir pointing cameras) are averaged on 4 by 4 pixel areas on-board the platform. This process achieves a 16:1 compression ratio in those data channels, and thus delivers data at the reduced spatial resolution of 1100 m. As a result, all standard MISR land products available from NASA are generated at this lower spatial resolution.

The MISR instrument can also be operated in Local Mode, upon request and at most once per orbit. In this Local Mode,
70   each camera is allowed to transmit data in turn at the native spatial resolution of the detectors in all 4 spectral bands, for a limited period of time, corresponding to a scene of about 300 km along-track. These data are useful for calibration purposes, as well as to provide information at the finer spatial resolution over particular sites, for instance during field campaigns.

For reference, during the initial pre-processing phase, the original analog radiance measurements acquired by the instrument's cameras are converted to 16-bit unsigned integers, where the first (most significant) 14 bits contain the scaled radiance
75   itself and the last 2 bits constitute the Radiometric Data Quality Indicator (RDQI), an indicator of the estimated reliability of the data item. The floating point radiance is obtained by unpacking those last 2 bits and multiplying the resulting value by the scale factor that is included in the data file. There is also a one-to-one relation between the radiance and the non-dimensional bidirectional reflectance factor (denoted Brf in this paper):

$$\mathrm{Brf} = [(\pi \times D^2)/(W \times \cos\theta_s)] \times \mathrm{Rad} \qquad (1)$$

80   where D is the 'approximate distance in astronomical units between the center of the Earth and the center of the Sun at the time MISR observes the first valid (non-fill) pixel in the swath. Replicated across bands for convenience; does not vary with band' (Bull et al., 2011, p. 73), $W = E_0^{std}(b)$ is the band specific weighted standardized solar irradiance in W m^{-2} μm^{-1} (Bull et al., 2011, p. 65–68), and Rad stands for the floating point radiance value, in W m^{-2} sr^{-1} μm^{-1}.

The on-board computer that manages the measurements and performs the Global Mode spatial averaging may occasionally
85   be overwhelmed by the high acquisition data rate, in particular—but not only—when switching from the default Global Mode to the Local Mode or conversely. When such an event occurs, the computer quickly resets itself to handle new data, and this may result in the loss of one or more line(s) of data in some or all of the 4 spectral data channels of one or more cameras.

Earth System
Science
Data
Open Access
Discussions
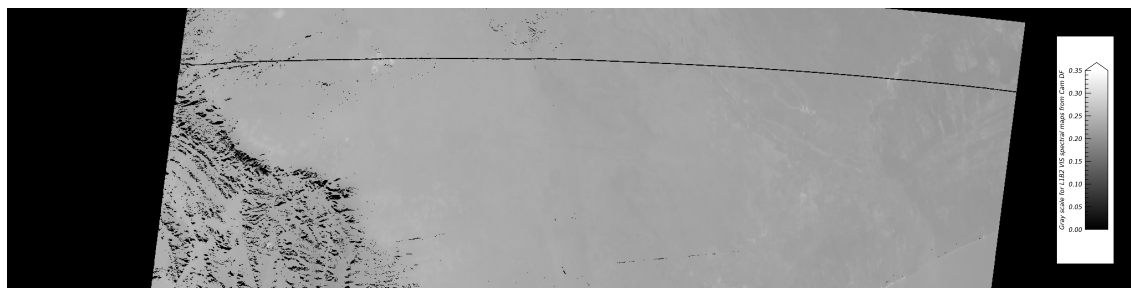
Figure 1. Example of missing lines (in this case, 2 consecutive lines) in the L1B2 terrain-projected Local Mode bidirectional reflectance factor data for the blue spectral band of the DF camera, for PATH 168, ORBIT 68050, and BLOCK 110, covering an area located in eastern South Africa. The linear grey scale, shown on the right side of the frame, assigns Brf values of 0 to black and Brf values of 0.35 or higher to white. The black spots on the western side of the map are due to pixels that are not observed at this slanted view angle due to the local topography. The total size of the Block area is 563.2 km across-track (roughly west-east) by 140.8 km along-track (roughly north-south), while parallelogram-shaped ground area inside the Block is about 380 km across-track.
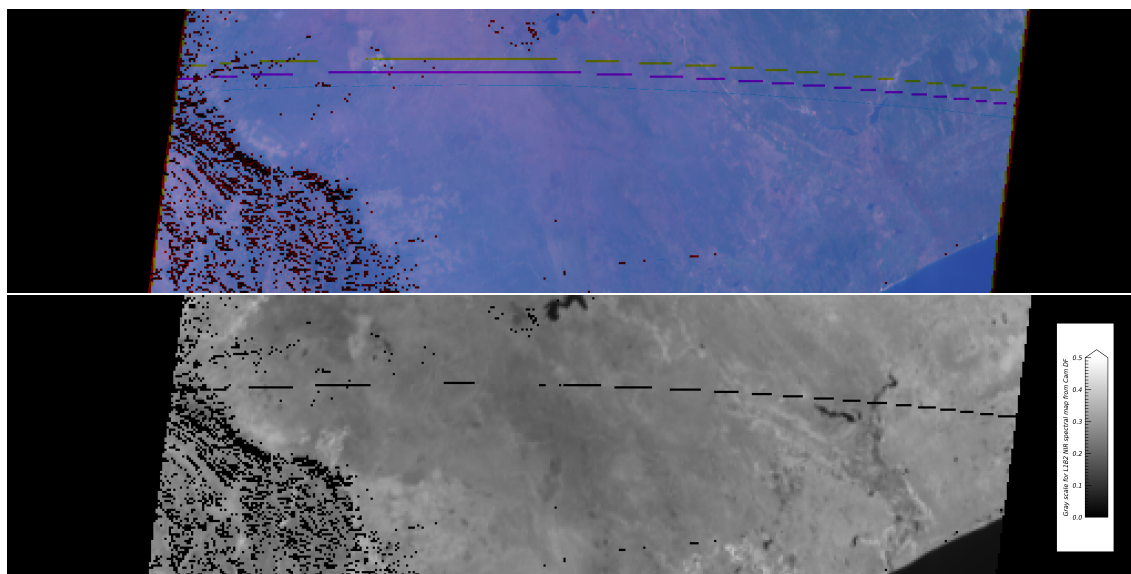
Figure 2. Top frame: Example of the effect of missing lines in a multispectral product, here an RGB composite of the red, green and blue spectral bands for the L1B2 terrain-projected Global Mode Brf data for the same DF camera case exhibited in Figure 1. The top (yellow), middle (magenta) and bottom (light cyan, barely visible because of the general bluish tone of the image) lines correspond to the missing data in the blue, green and red spectral bands, respectively. This image was generated by first replicating the low resolution blue and green pixels so as to match the resolution of the red band, and then creating the composite image at the full spatial resolution. Bottom frame: Map of the NIR Brf of the same DF camera. Linear dimensions are the same as in Figure 1. The 3 RGB spectral channels are stretched from 0.0 to 0.35 for the purpose of generating the color image, while the NIR data is stretched between 0.0 and 0.5.

Figure 1 exhibits a simple example: in this case, a couple of lines have been dropped from the blue spectral band in BLOCK 110 of the DF camera in the Local Mode bidirectional reflectance factor data file. In this context, a BLOCK is a segment of data that covers a ground area of about 563.2 km across-track by 140.8 km along-track, which contains usable data over roughly 380 km across-track, as well as fill data on both edges. These missing lines tend to occur simultaneously in all spectral bands, but because the spectral detectors are located next to each other on the focal plane of each camera, the corresponding locations of these missing values on the ground, and therefore in the images, are slightly displaced. As a result, products that combine multiple spectral bands from the same camera inherit all missing lines from each of the individual bands. And of course, a missing line in a Local Mode data also implies a missing line in the corresponding Global Mode data, with an added side effect: the missing lines in the non-red spectral channels of the off-nadir cameras show up with a width of 1100 m (or multiples thereof). Figure 2 exhibits the RGB image for the Global Mode acquisition corresponding to the case shown in Figure 1 (top frame), as well as the corresponding map of the NIR spectral band for the same camera.

2.2 Detecting missing values

The first step in this investigation consists of documenting the prevalence of this problem. As explained above, radiance values are represented in the L1B2 GRP data files by 16-bit unsigned integers, where the last 2 bits are the RDQI. Table 1 describes the meaning attached to the values of this quality indicator (Bull et al., 2011, p. 71).

Table 1. Meaning of the 2-bit RDQI values in L1B2 GRP data files.

Decimal	Binary	Meaning
0	00	Within specifications ('good')
1	01	Reduced accuracy ('fair')
2	10	Not usable for science ('poor')
3	11	Unusable for any purpose ('missing')

An RDQI value of 3 (binary 11) can be assigned to a terrain-projected L1B2 GRP value in 4 different cases, and the radiance value itself indicates why those values are unusable:

1. when the location of a pixel on the ground is obscured by topography for the specified camera, an issue that mostly affects off-nadir pointing cameras: its value is coded as the unsigned 16-bit integer 65511, equivalent to a scaled 14-bit radiance of 16377,

2. when the location of a pixel on the ground lies outside of the swath width of the camera, which occurs on the western and eastern edge of each BLOCK: its value is coded as the unsigned 16-bit integer 65515, equivalent to a scaled 14-bit radiance of 16378,

3. when the location of a pixel at the Earth's surface belongs to a BLOCK covering an ocean region only: its value is coded as the unsigned 16-bit integer 65519, equivalent to a scaled 14-bit radiance of 16379, or

4. when the radiance estimated for a location is deemed unusable for a reason other than orography, edge or ocean: its value is coded as the unsigned 16-bit integer 65523, equivalent to a scaled 14-bit radiance of 16380: this is a 'missing' value
115 in this context.

2.3 Counting missing and poor values

An examination of the MISR L1B2 GRP Terrain-Projected Global Mode (GM) Radiance files for BLOCKs 110 to 113 of PATHs 168 to 170, from the start of the mission onwards, revealed that millions of pixel values went missing (RDQI of 3 and scaled radiance of 16380) in Global as well as Local Mode acquisitions. For instance, Figure 3 exhibits the logarithm in base
120 10 of the number of missing pixels in the 4 spectral bands of the 9 cameras for BLOCK 110 of all available ORBITs of PATH 168 between March 24, 2000 and December 23, 2018: A total of 2,143,033 pixel values went missing. Local Mode data were briefly acquired at the start of the MISR mission for the Skukuza flux tower site located in the Kruger National Park of South Africa, in support of the Safari-2000 campaign. This site, which covers an area roughly equivalent to BLOCKs 110 and 111, has subsequently been systematically acquired from December 2009 onward. The net effect of switching between Global to Local
125 Mode (and back) on the number of missing values is clearly visible in this logarithmic plot: the number of missing values is about an order of magnitude larger when Local Mode acquisitions are taking place.
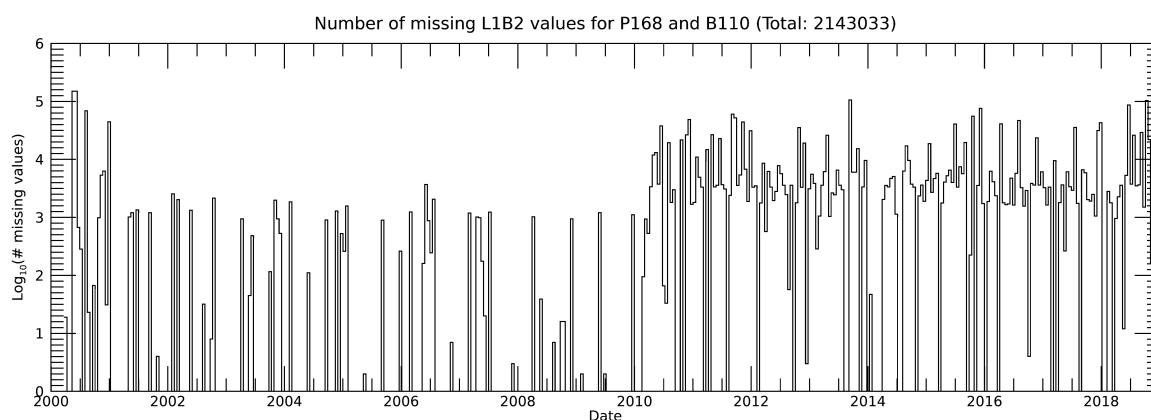
Figure 3. Time series of the logarithm in base 10 of the number of missing pixel values, in the 4 spectral bands of the 9 cameras of BLOCK 110 from each available MISR L1B2 GRP ORBIT in PATH 168, between March 2000 and December 2018. This BLOCK includes the Skukuza flux tower located in the Kruger National Park of South Africa. Local Mode data were briefly acquired at the start of the mission to support the Safari-2000 campaign, and then systematically from December 2009 onward. It can be seen that there is a strong relation between the number of missing values and the acquisition of Local Mode data, although the latter is not the only process responsible for missing data. See text for details.

A close inspection of the data quality in these files also revealed that missing values are often associated with 'poor' or 'fair' radiance values (with an RDQI value of 2 or 1, respectively) in their immediate vicinity. Figure 4 shows the time series of all poor values found in the 4 spectral bands and 9 cameras for BLOCK 110 of all available ORBITs of PATH 168 between the

130   start of the mission and December 23, 2018. Some 918,420 values fell in that category during that period: that brings the total number of missing or poor values to over 3 million for that particular PATH and BLOCK.
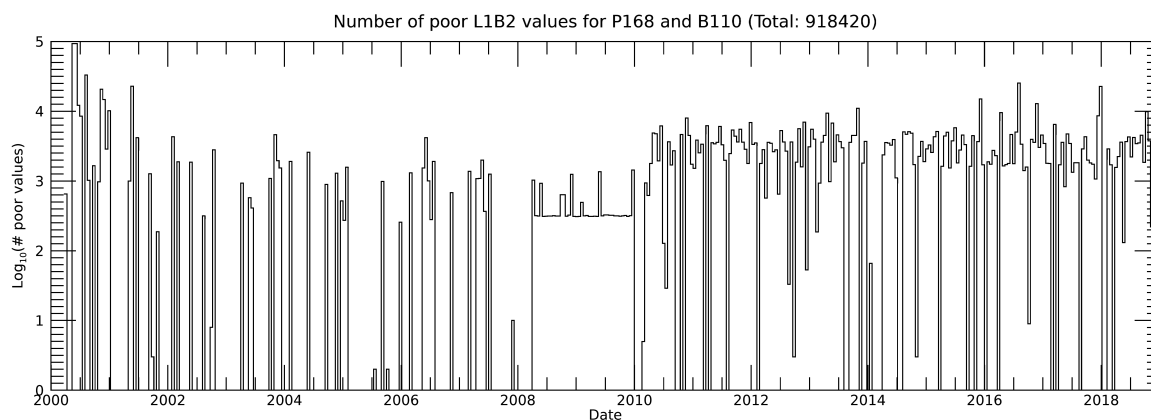
Figure 4. Time series of the logarithm in base 10 of the number of poor pixel values, in this case in the 4 spectral bands of the 9 cameras of BLOCK 110 from each available MISR L1B2 GRP ORBIT in PATH 168, between late March 2000 and December 2018 (the same BLOCK as in Figure 3). See the text for a discussion of the similarities and differences between those two time series.

These findings call for multiple comments:

– While these numbers appear large (out of context), they remain tiny compared to the total number of 'good' and 'fair' pixel values (i.e., those associated with a RDQI of 0 or 1), which is of the order of 25×10^9 pixels for the same BLOCK

135   and time period. So this problem does not affect the usability of the MISR instrument, especially for studies involving large areas or long periods of time.

– However, for the technical reasons described above, missing values are often concentrated in geographical areas where Local Mode acquisitions take place. Hence, if a site is systematically acquired in Local Mode over a substantial period of time, then that region tends to be more affected by this problem than other areas. This, in turn, implies that local studies,

140   field campaigns or projects that occur precisely in those areas and time periods may be severely impacted while similar investigations elsewhere may be relatively unaffected.

– There are somewhat fewer poor values than missing values in general, but there is also less difference in the peak number of poor values with or without Local Mode acquisitions. This is because measurements affected by sun glint are typically marked as poor. This natural phenomenon occurs systematically and may affect a camera irrespective of the Local or

145   Global Mode of acquisition, especially whenever the observed scene contains substantial water bodies.

  – Nevertheless, as was the case for L1B2 radiance values, the number of poor values also increased by about an order of magnitude whenever the Skukuza site was systematically observed in Local Mode.

  – Figure 4 also reveals that a significant number of radiance measurements were assigned an RDQI of 2 between about March 2008 and the end of 2009, for reasons that remain to be investigated but that probably involve changes in the software processing system, rather than the performance of the instrument.

The study of Mahlangu et al. (2018) provides a good example of the crippling effect of these missing and poor data on a particular investigation: An airborne LiDAR field campaign took place in and around the Skukuza site in April 2012. The aim of that study was (1) to document the structure of vegetation using the LiDAR as a high spatial resolution reference data over a limited area, and (2) to explore whether a relation between these airborne measurements and MISR-derived products could help describe those ecosystem properties over a larger region.

This study relied on products to characterize the plant canopy generated by the MISR-HR processing system (Verstraete et al. (2012), Version 1.04-5) for the closest available acquisition, namely PATH 168, ORBIT 65487 and BLOCK 110, acquired on 10 April 2012. Figure 5, demonstrates the problem: the spatial coverage of the Leaf Area Index (LAI) product over terrestrial areas was significantly reduced by clouds (screened out on square areas of 17.6 km by 17.6 km by the software and data available at the time that paper was generated) and the remaining clear area was very much contaminated by missing data lines in all 4 spectral bands of 2 different cameras (AF and CA).
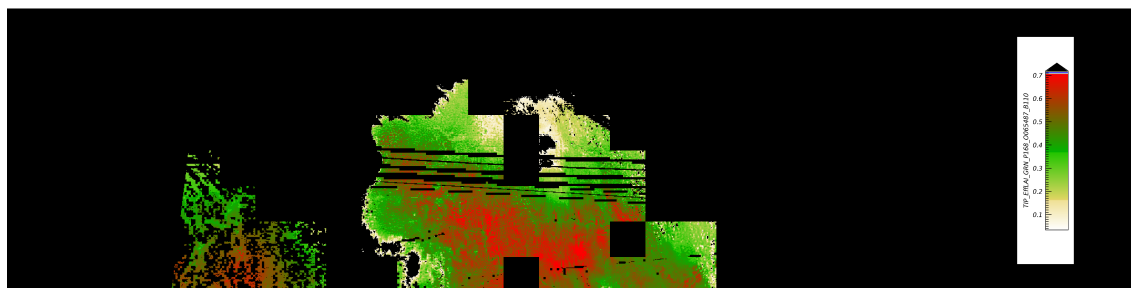
Figure 5. Example of the effect of missing lines in a derived product, here the Leaf Area Index (LAI) for the particular date and time when an airborne field campaign was taking place in the Kruger National Park of South Africa (PATH 168, ORBIT 65487, BLOCK 110; see text for details). A large fraction of the MISR BLOCK is unavailable (depicted in black), due to the prevailing cloud cover on that particular date as well as the presence of missing lines. The color gradient (white, yellow, green, red) maps the LAI range [0.0, 0.7].

To address this issue and avoid serious contamination of subsequent analyses and products generated by the MISR-HR processing system (Verstraete et al., 2012), the radiance values that are missing, and optionally those associated with an RDQI of 2, should be replaced by best estimates of what the original measurements should have been. The next sections describe the proposed solution to this problem and document its performance.

3 Materials and Methods

3.1 MISR data

The MISR mission is run by a group of engineers and scientists at NASA's Jet Propulsion Laboratory (JPL) in Pasadena,
California, supported by an international team of researchers, while the data generated by this instrument and the standard
170 products derived from the raw measurements are processed and permanently archived at NASA's Atmospheric Science Data
Center (ASDC) in Hampton, Virginia. As is the case for other Earth Observation missions, all MISR data products are freely
available to anyone interested, either through customized orders, or through FTP access to the Data Pool at
`https://eosweb.larc.nasa.gov/project/misr/misr_table`.

Three standard data sets are of direct relevance for the current purpose: the AGP and two out of the 6 parameter sets of
175 the Level 1B2 Georectified Radiance Product (GRP): the terrain-projected Top of Atmosphere (ToA) radiance Global Mode
parameter and the Radiometric Camera-by-Camera Cloud Mask (RCCM) (Bull et al., 2011, p. 56).

3.1.1 The AGP dataset

The Ancillary Geographic Product (AGP) is the master reference dataset containing essential information on the latitude and
longitude of BLOCK locations, as well as ancillary data such as the expected distribution of land masses and water bodies (in
180 particular oceans) (Bull et al., 2011, p. 14). This primitive land cover map (Version `F01_24` at the time of this writing) actually
recognizes 7 different types of "surface features" (Bull et al., 2011, p. 210):

– 0: Shallow Ocean

– 1: Land

– 2: Coastline

185 – 3: Shallow Inland Water

– 4: Ephemeral Water

– 5: Deep Inland Water

– 6: Deep Ocean

The spatial distribution of these feature types is fixed throughout the mission: it is not intended as a genuine, detailed land
190 cover map, but as a coarse land-sea partition, with a few more features. In the current context, categories 0, 5 and 6 are merged
into a water class, and the remaining categories are assigned to the land class.

3.1.2 The Terrain-projected ToA radiance Global Mode parameter set

This parameter set is the primary MISR radiance product at the nominal Top of the Atmosphere (ToA): all higher level
products, including the RCCM described in the next subsection, are derived from an analysis of those calibrated and geolocated

195   measurements. In the default Global Mode of operation, two versions of that parameter set are generated: one in which all measurements are projected on the World Geodetic System 1984 (WGS84) reference ellipsoid, primarily used in applications over oceans, and one where measurements are projected on a Digital Elevation Model (DEM), primarily used in terrestrial applications. This paper is concerned exclusively with this latter projection. All results described in this paper are based on L1B2 Georectified Radiance Product Version `F03_0024`.

3.1.3   The RCCM parameter set

200   The Radiometric Camera-by-camera Cloud Mask (RCCM) provides a separate cloud mask for each of the 9 MISR cameras. This standard data set is described in detail in the Level 1 Cloud Detection Algorithm Theoretical Basis Document (ATBD) (Diner et al., 1999b), which can be downloaded from `https://eospso.gsfc.nasa.gov/atbd-category/45`, and in subsequent papers such as Zhao and Di Girolamo (2004) and Yang et al. (2007). All results described in this paper are based

205   on RCCM Version `F04_0025`.

Over land, these camera-specific cloud masks are derived from the L1B2 radiance data described in the previous subsection, using the red and the near-infrared (NIR) spectral bands of each camera. Hence, whenever there are missing data in those spectral bands, there will be missing data in the RCCM product too. However, a procedure to replace missing data in those cloud masks has recently been designed and is fully documented in Verstraete et al. (2019).

3.2   Correlation between the target and the sources

210   A practical solution to the problem described in Section 2 above (i.e., restoring data when there are missing lines) is made possible by two findings: (1) as shown earlier, the missing pixels appear in different locations in different spectral channels of the affected camera, and (2) in most cases, it is possible to find, for each data channel affected by missing or poor data, one or more data channels (among the 35 others simultaneously acquired, i.e., for the same MISR PATH, ORBIT and BLOCK) that

215   exhibit a very high correlation with the one in need of an update. In the rest of this paper, the data channel that is being repaired is called the *target*, and the other 35 data channels of the same data acquisition are called the *sources* (or predictors).

The Pearson correlation coefficients (CC) between all pixel values that are valid (i.e., excluding all values with an RDQI value of 2 or 3) and common between the target and each of the 35 source data channels are computed. In each of these 35 cases, the root mean square difference (RMSD) between the data channel values, the best linear fit between the source and the

220   target, and the χ^2 value, i.e., a measure of the dispersion of the data points around that linear fit function, are also calculated. This set of statistics is then ranked in decreasing order of correlation coefficient value. It turns out that one or more source data channels are usually highly correlated (CC larger than 0.9) with the target data channel, thereby providing reliable means to replace the missing values in the target on the basis of non-missing values in the most promising source data channel.

This approach is first demonstrated on the MISR GRP Global Mode data for PATH 168, ORBIT 68050, BLOCK 110, camera

225   DF exhibited in Figure 2. Table 2 shows the statistics describing the relations between the 4 target spectral bands of the DF camera and the best (i.e., the most promising, in terms of the Pearson CC) identified sources (in this case the same spectral channels of the CF camera), using all valid pixels that are common to both data channels. The χ^2 value for the red band (or

indeed for any high resolution data channel) is much larger than for the other Global Mode data channels because they contain about 16 times more data points.

Table 2. Correlation statistics between the radiance in the 4 spectral bands of the target (Tgt) camera DF and the radiance in the best correlated source (Src) data channels (first attempt), using all common and valid values for the GLOBAL MODE measurements acquired in PATH 168, ORBIT 68050, and BLOCK 110. The columns entitled 'No. Pix.', 'RMSD', 'PCC' and 'χ^2' provide the number of valid measurements common to both data channels, the root mean square difference between the data channel values, the Pearson correlation coefficient between them and the value of that statistics for the linear fit between those data channels, respectively.

| Tgt. Cam.	Tgt. Band	Src. Cam.	Src. Band	No. Pix.	RMSD	PCC	χ^2
DF	Blue	CF	Blue	45779	16.32	0.951	71875
DF	Green	CF	Green	45941	5.283	0.969	53560
DF	Red	CF	Red	762392	2.409	0.977	2.23E+06
DF	NIR	CF	NIR	46186	1.741	0.989	129323

230   Figure 6 shows the scatterplots and the linear fits between the 4 target spectral bands of the DF camera case mentioned in Table 2 and the best predictor source data channels for the selected BLOCK.

While the correlations are highly significant, in part due to the large number of pixel values, the shape and the dispersion of the cloud of data points often suggests that a single linear equation may not always be optimal to predict the target values across the entire Block. This is largely due to the fact that the scene may include rather different geophysical media, which

235   follow possibly different statistical relations as far as the correlations between data channels are concerned.

To explore this avenue, the sets of points used in these computations were partitioned in 3 categories: clear land masses, clear water bodies, and cloud fields. The geographical distribution of land masses and water bodies was extracted from the MISR surface cover type map contained in the static AGP file described earlier, for the specified PATH and BLOCK. The cloud information was retrieved from the Radiometric Camera-by-Camera Cloud Mask (RCCM) data product, updated as

240   explained in Verstraete et al. (2019). Statistics, including the Pearson correlation coefficients and the best linear fit equations, were generated for each of these 3 categories. The gain in pursuing this approach is demonstrated in Figure 7, which compares the best fit statistics while processing the NIR spectral band of Camera AA for the same scene as above, when all valid values available in the BLOCK are used for establishing the statistics (left panel), and when only those values attributable to clear land masses are used (right panel). The RMSD, the CC and the χ^2 statistics are significantly improved by focusing on the clear land
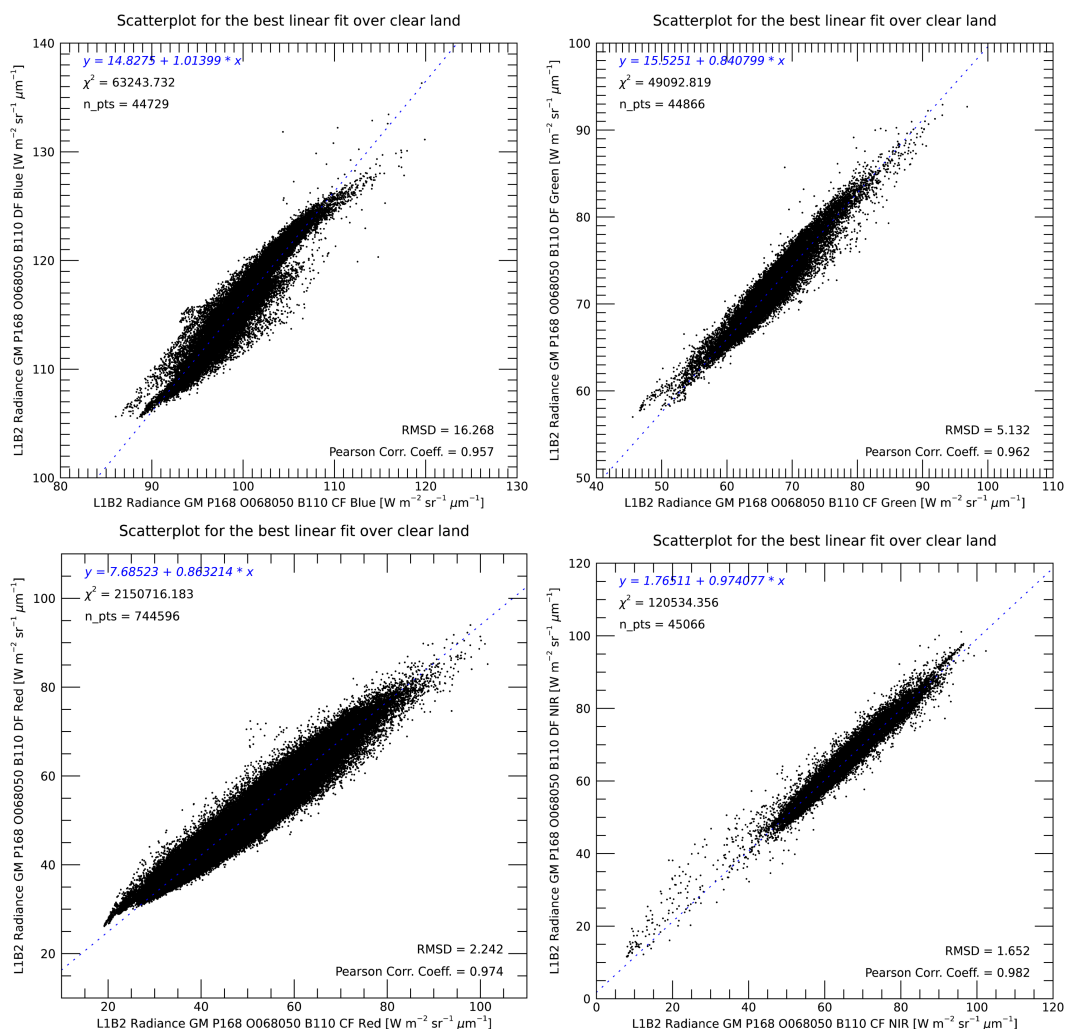
245   masses only.
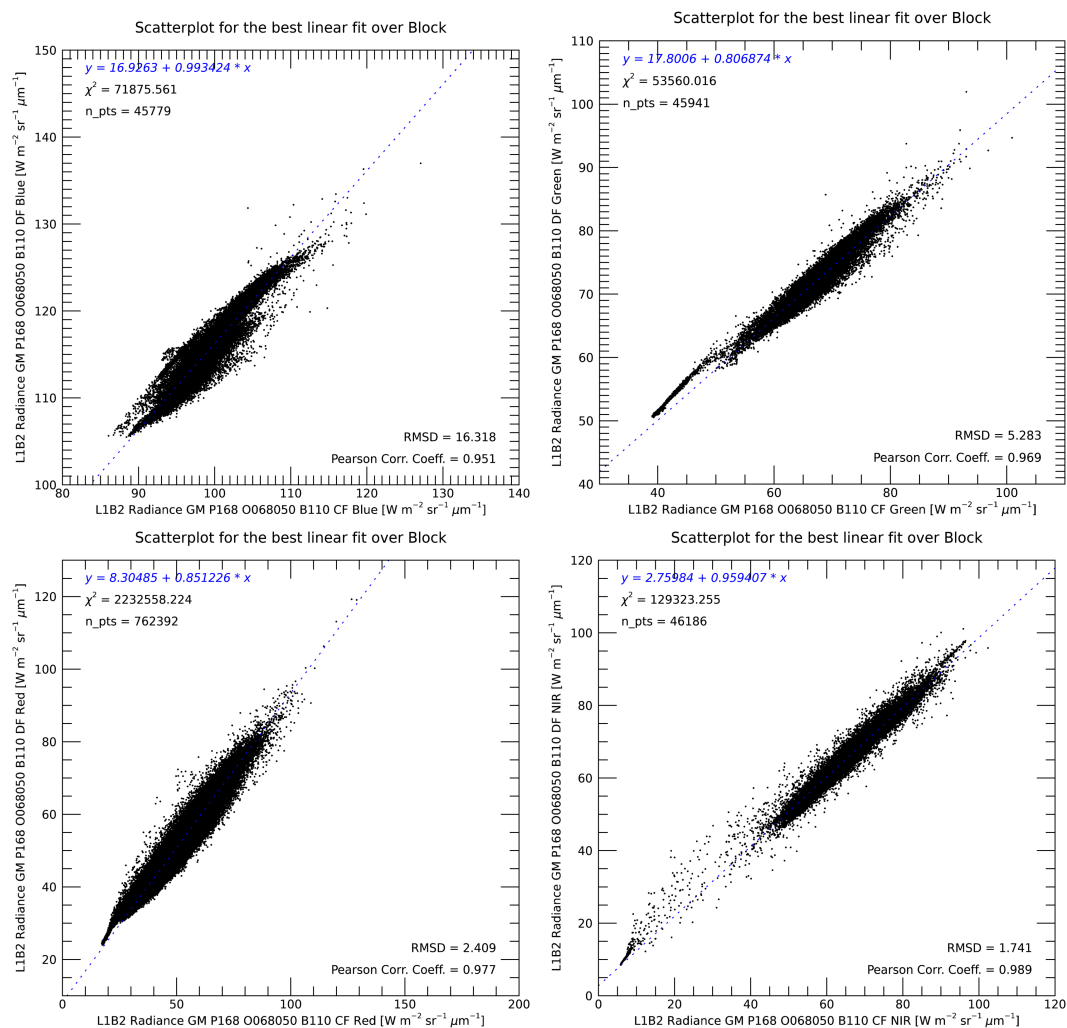
Figure 6. Scatterplots and best linear fit relations to predict the radiance values of the 4 spectral bands of the target DF camera for the GLOBAL MODE measurements of PATH 168, ORBIT 68050, and BLOCK 110, using all common valid values. In this case, the best predictors for the measurements by the DF camera were those acquired by the CF camera in the same spectral bands as the targets (Top left panel: blue, top right panel: green, bottom left panel: red, and bottom right panel: NIR). Note that the ranges of values are different for each plot in order to optimize the view of the data.
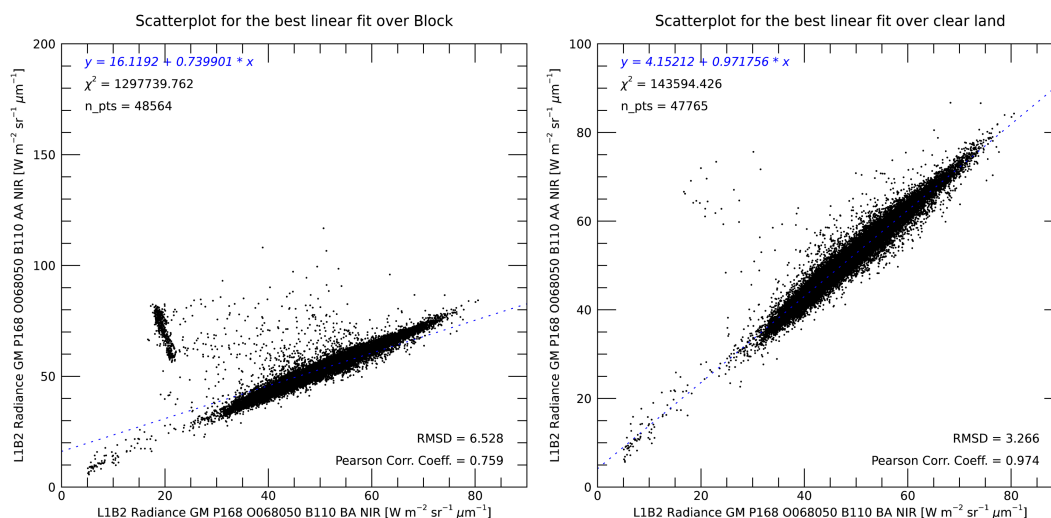
Figure 7. Scatterplots and best linear fit relations to predict the radiance values of missing data in the NIR spectral band of the AA camera, for the same GLOBAL MODE measurements of PATH 168, ORBIT 68050, and BLOCK 110 as in Figure 6. The left panel shows the results obtained when all common valid values throughout the BLOCK are considered, while the right panel shows the statistical results applicable only to the clear land masses. The Pearson correlation coefficient, the RMSD and the χ^2 statistics are definitely improved in this latter case, as a group of more than 800 pixel values, identified as clouds, is treated separately.

Table 3 shows the statistical results for the same cases as before (Table 2), this time reporting values for clear land masses only. While the correlation coefficients are roughly equivalent, despite the somewhat smaller number of points involved, the dispersion statistics (RMSD and χ^2) are again improved. Figure 8 exhibits the corresponding scatterplots and the best linear fit functions for clear land masses: as expected, these compare favorably to the results shown in Figure 6, since this particular
250   BLOCK is essentially cloud-free.

Table 3. Correlation statistics between the 4 spectral bands of the target camera DF and the best correlated source data channels (first attempt), using all common and valid clear land values for the GLOBAL MODE measurements acquired in PATH 168, ORBIT 68050, and BLOCK 110. The column headers are identical to those in Table 2.

Tgt. Cam.	Tgt. Band	Src. Cam.	Src. Band	No. Pix.	RMSD	PCC	χ^2
DF	Blue	CF	Blue	44729	16.27	0.957	63244
DF	Green	CF	Green	44866	5.132	0.962	49093
DF	Red	CF	Red	744596	2.242	0.974	2.15E+06
DF	NIR	CF	NIR	45066	1.652	0.982	120534

Figure 8. Scatterplots and best linear fit relations to predict the clear land surface values of the 4 spectral bands of the DF camera for the GLOBAL MODE measurements of PATH 168, ORBIT 68050, and BLOCK 110. In this case, the best predictors for the measurements by the DF camera were those acquired by the CF camera (Top left panel: blue, top right panel: green, bottom left panel: red, and bottom right panel: NIR). Note that the ranges of values are different for each plot in order to optimize the view of the data.

3.3 Replacement algorithm

These findings suggest that missing and optionally poor L1B2 radiance values could be replaced by reconstructed estimates based on the actual measurements obtained quasi simultaneously in other data channels (spectral bands or cameras), for the same geographical location. This Section outlines the sequence of steps undertaken to achieve this goal, in a schematic manner intended to also serve as a guide to the IDL function `fix_l1b2.pro` which implements these concepts in the operational

255

14

processing software. Unless noted otherwise, all results shown below are derived using Version 2.1.x or later of the IDL software available on GitHub.

This function takes as input arguments pointers to the scaled radiances (with the RDQI attached), the unscaled radiances, the Brf and the RDQI data buffers for the specified MISR MODE, PATH, ORBIT and BLOCK, which must have been loaded
260 on the heap prior to the call. This ensures a fast access to those data fields, as they will need to be read multiple times during the processing. An initial suite of preliminary steps is then undertaken to set the stage:

- The IDL function offers the option of mapping the L1B2 product before processing, to be able to show how the results of the processing differ from the original MISR data.

- The next step consists in generating the masks for clear land masses, clear water bodies and cloud fields, to partition
265 all valid radiance values found in the BLOCK into those 3 categories. Although statistics are in fact computed for clear water bodies and cloud fields, they are not reported here, as the land areas are of primary interest.

- An optional main log file may be created to keep track of the progress and provide information on intermediary results, for instance for diagnostic purposes.

- The function `pre_fix_l1b2.pro` is then called to locate, within each of the 36 data channels, the numbers and
270 positions of the poor and missing data values. If any are found, the function `best_fits_l1b2.pro` is activated to compute the optimal statistics and the best linear fit functions to replace those poor or missing values, and to optionally generate a separate log file for each case. Information on progress is also added to the main log file if it has been requested in the previous step.

At that point, the code relies on 2 closely related functions, `fix_poor_l1b2.pro` and `fix_miss_l1b2.pro`, to
275 replace poor (if required) and missing L1B2 values, respectively. The core of the processing is outlined in the following pseudo-code listing:

```
FOR each camera
  FOR each spectral band
    access the MISR L1B2 data for the target
280  retrieve the number and location of poor or missing values
    skip to the next camera/band combination if no values need replacement
    WHILE the maximum number of attempts has not been reached
      IF poor or missing values remain in the land/water/cloud target
        record the current best source and correlation statistics
285      access the MISR L1B2 data for the current source
        upscale or downscale the source data to match the target resolution
        FOR each target value that needs replacement
          record the radiance value at the same location in the current source
```

15

Earth System
Science
Data

Open Access

Discussions

```
                 ensure that this source value is valid
290              apply the best linear fit equation for this attempt
                 replace the poor or missing value in the target by this new estimate
                 update the RDQI for that new value to 1B
             ENDFOR
             update the number and location of remaining missing (or poor) values
295      ENDIF
         point to the next best source and correlation statistics
       ENDWHILE
     ENDFOR
   ENDFOR
```

300 If poor values are replaced, the function `fix_poor_l1b2.pro` also optionally generates statistics and scatterplots of the updated versus the original values.

The `WHILE` loop over `attempts` is required because in some cases the best source data channel to replace values in the target data channel happens to also have missing values in the very same location as the target. In those cases, the next best data channel must be considered. The maximum number of attempts (i.e., the maximum number of best source data channels 305 to consider) is specified as an argument to the function. Practical experience and considerations indicate that this value should generally be set to no more than `4`: smaller values may result in non-trivial numbers of missing or poor values not being replaced, while larger values may result in the use of source data channels that are less well correlated to the target, and therefore result in proposing less reliable replacement values. However, see Subsection 4.3 below for an exceptional case that required 8 iterations to replace most missing values. Note also that replaced values are always assigned an RDQI of `1`.

310 ## 4 Results

The following exhibits will show graphically the outcome of this process, for the cases discussed in Section 2 above.

4.1 Global Mode, Path 168, Orbit 68050, Block 110, Camera DF

The particular case exhibited earlier in Figure 2, where 4 spectral bands of the DF camera exhibit missing and poor data values, is examined first.

315 Figure 9 shows the outcome of the process where each and every missing value in the original target data channels is replaced by a value estimated by retrieving the valid bidirectional reflectance factor values retrieved from the source data channels indicated in Table 3 for the very same geographical location, and inserted into the linear fit equations exhibited in Figure 8. The top frame is an RGB composite and the bottom frame shows the map of the corresponding NIR spectral band. These frames can be compared to those exhibited in Figure 2.
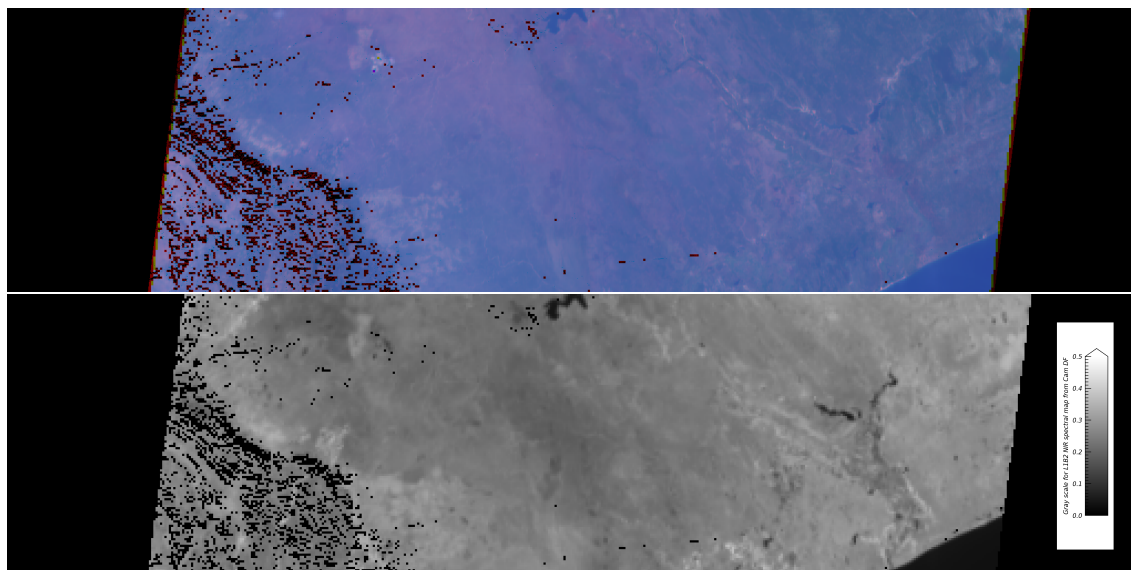
Figure 9. Top frame: RGB composite map of the red, green and blue spectral bands for the L1B2 terrain-projected Global Mode bidirectional reflectance factor data file for the same DF camera case exhibited in Figure 2. Bottom frame: Map of the NIR spectral band for the same DF camera. All missing values in the original data files have been replaced by estimates as described in the text. Linear dimensions are the same as in Figure 1, and all 3 spectral channels are stretched from 0.0 to 0.35 for the purpose of generating the color image; the gray scale used to stretch the NIR data for the bottom frame assigns black to reflectance values of 0.0 and white to values of 0.50 or higher.

4.2 Global Mode, Path 168, Orbit 65487, Block 110, Camera CA

The case exhibited earlier in Figure 5 is examined next. Table 4 summarizes the correlation statistics between the 4 spectral bands of the CA camera, all affected by missing values, and the best sources to estimate replacement values, considering only the valid values in clear land masses that are common to the target and the source data channels.

Table 4. Correlation statistics between the 4 spectral bands of the target camera CA and the best correlated source data channels, using only common and valid clear land values for the GLOBAL MODE measurements acquired in PATH 168, ORBIT 65487, and BLOCK 110. The column headers are identical to those in Table 2.

Tgt. Cam.	Tgt. Band	Src. Cam.	Src. Band	No. Pix.	RMSD	PCC	χ^2
CA	Blue	CA	Green	17591	22.48	0.966	156441
CA	Green	CA	Blue	17591	22.48	0.966	119634
CA	Red	DA	Red	268707	10.47	0.955	1.14E+07
CA	NIR	DA	NIR	15613	6.866	0.969	58113

17

Figure 10 exhibits maps of the RDQI for those 4 data channels of the CA camera. The relative displacement of missing lines
325   between the spectral bands and the clustering of poor values around those lines are clearly observable: these findings have been
widely confirmed in all cases investigated.
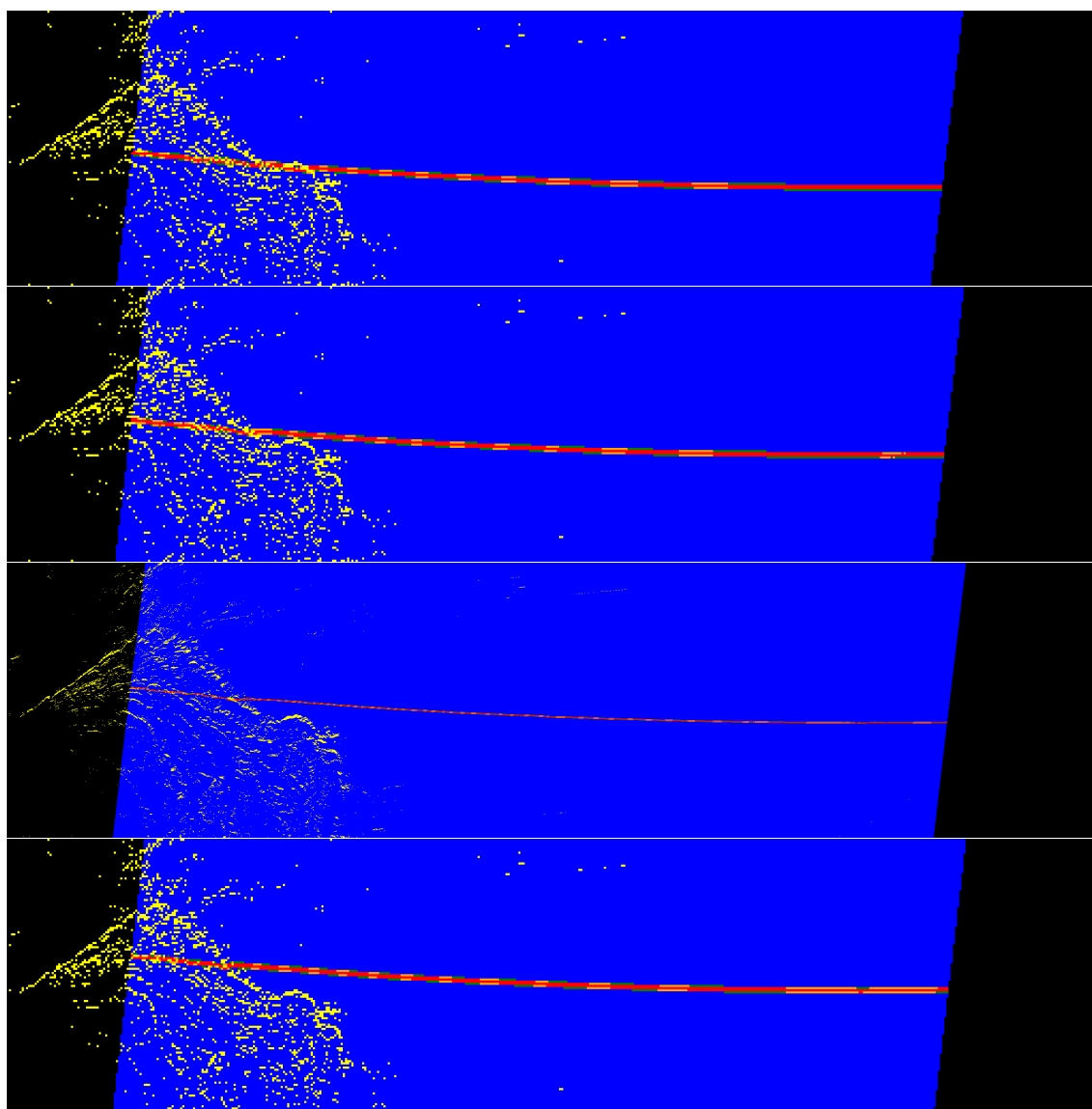
Figure 10. Spatial distribution of RDQI values for GLOBAL MODE, PATH 168, ORBIT 65487, BLOCK 110, CAMERA CA, in the 4 spectral
bands: Blue, Green, Red, and NIR from top to bottom. Color coding: Black: Edge pixels; Yellow: pixels obscured by the local topography;
blue: RDQI = 0, green: RDQI = 1; orange: RDQI = 2; red: RDQI = 3.

Note that the best predictor for missing and poor values in the CA/Blue data channel is the CA/Green data channel, and
conversely. Since the RMSD and the Pearson correlation coefficient are symmetric with respect to their inputs, the statistics are

identical in these cases. However, the CA/Green data points fit a little bit better the best linear function (the χ^2 for the function
330 predicting CA/Green is somewhat lower than that for CA/Blue). The best predictors for the CA/Red and CA/NIR data channels
are provided by the DA camera, in the corresponding spectral bands.

The scatterplots for these 4 cases are exhibited in Figure 11. It can be seen that the best fit function to predict the CA/Red
data channel is not as good as for the other data channels, although the Pearson correlation coefficient is still 0.836. These
statistics are of course also affected by the fact that there are quite a lot fewer valid data values available to estimate the missing
335 and poor pixels, compared to the earlier case of ORBIT 68050.
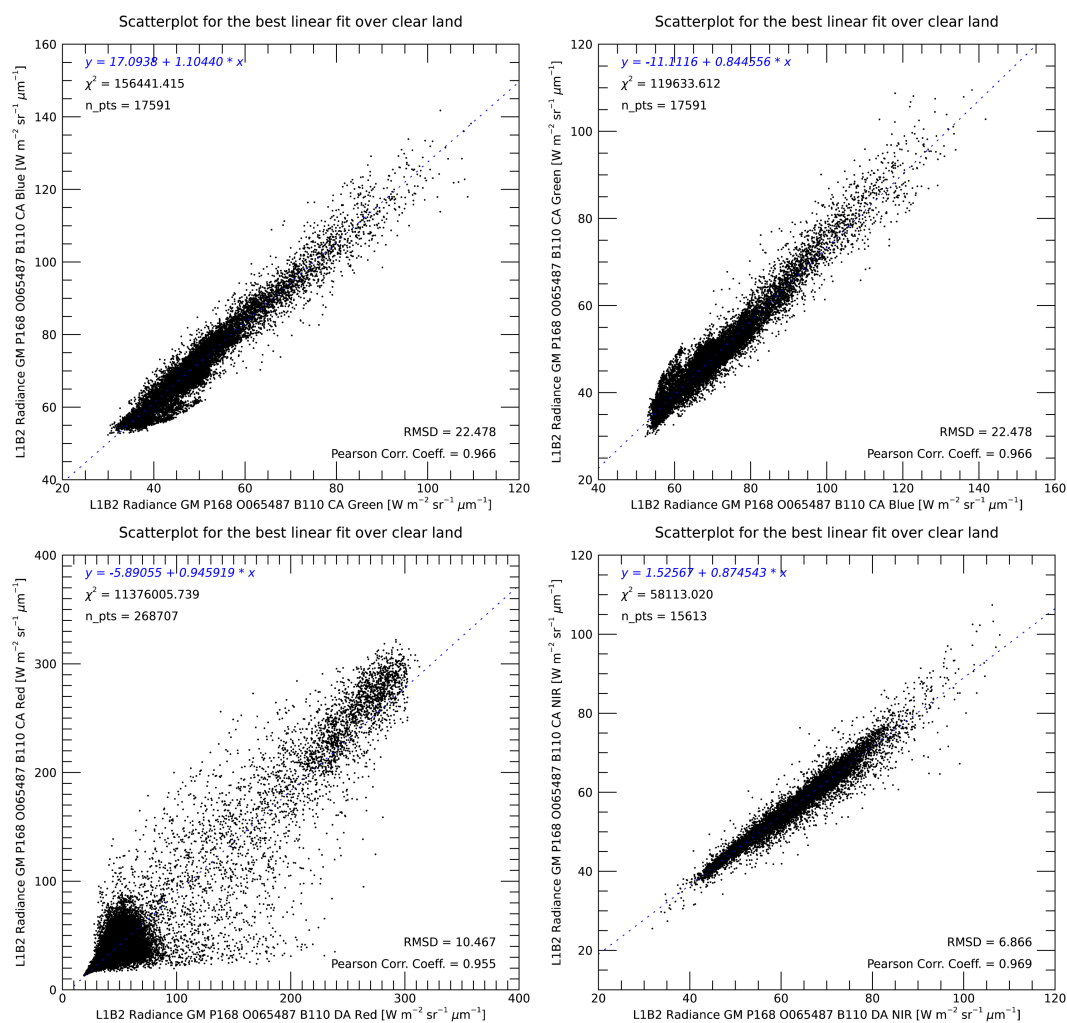
Figure 11. Scatterplots and best linear fit relations to predict the values of the 4 spectral bands of the target CA camera for the GLOBAL
MODE measurements of PATH 168, ORBIT 65487, and BLOCK 110. Top left panel: Blue, top right panel: Green, bottom left panel: Red, and
bottom right panel: NIR.

19

Using those best linear functions and the valid values in the source data channel, it is possible to generate estimates of the missing and poor values in the target data channel. Figures 12 and 13 exhibit the RGB composite map of the first 3 spectral bands and the B&W map of the NIR spectral band of the CA camera, before and after the replacement of missing values.
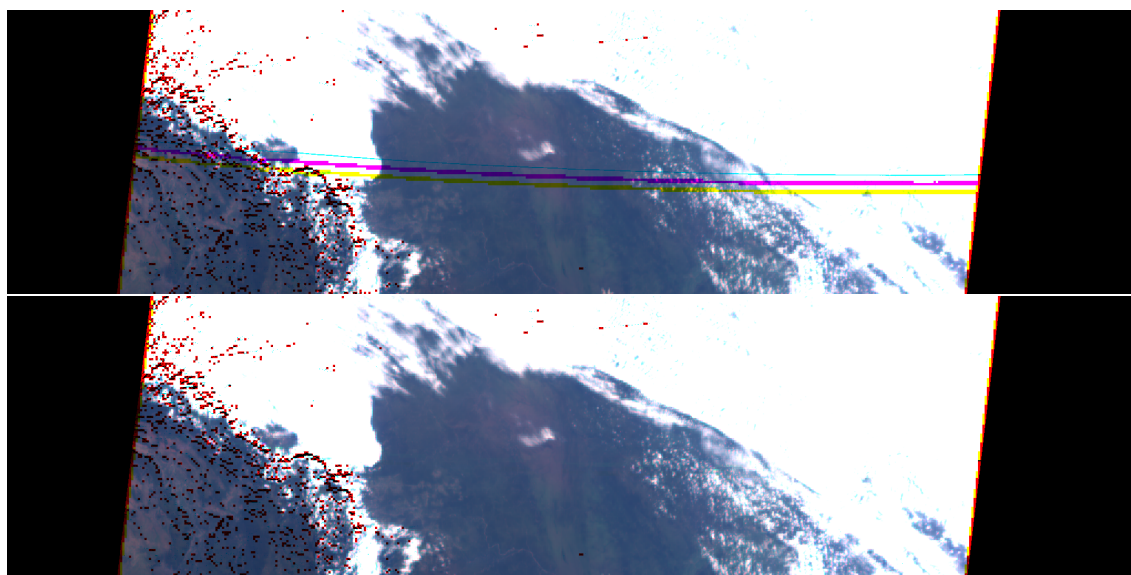
Figure 12. RGB maps of the 3 visible data channels for GM-P168-O065487-B110-CA before (top) and after (bottom) replacing the missing and poor data values by the best estimates computed as described above.
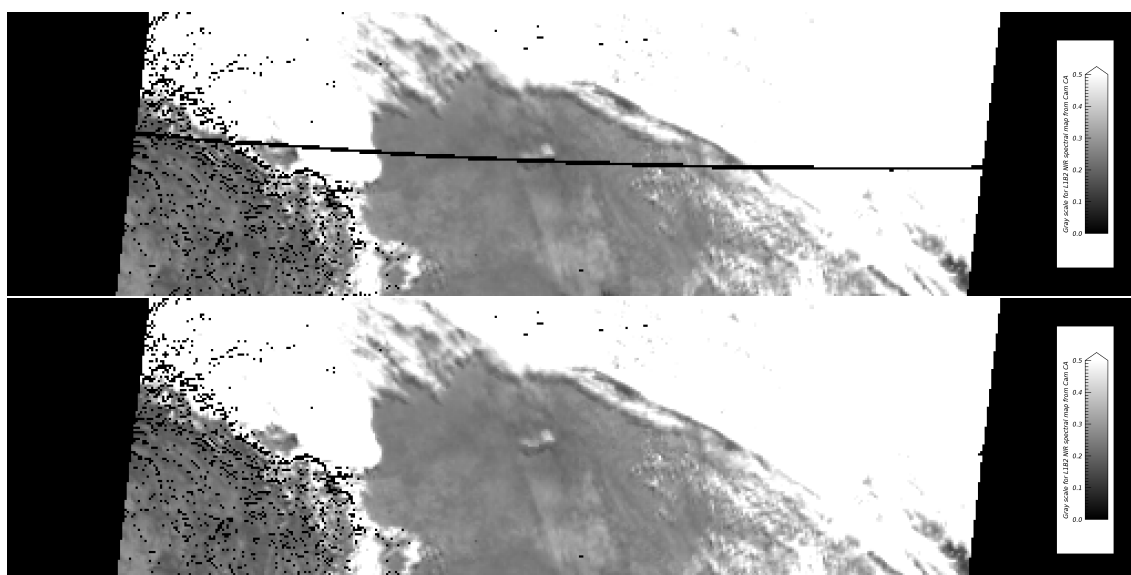
Figure 13. Maps of the GM-P168-O065487-B110-CA-NIR Brf data channel before (top) and after (bottom) replacing the missing and poor data values by the best estimates computed as described above.

4.3   Global Mode, Path 168, Orbit 2111, Block 110, Camera BA

340   The next instance exhibits the most severe case of missing values encountered so far in our investigations. Figure 14 shows the RDQI values for the 4 spectral bands of the BA camera for P168, O002111 and B110.
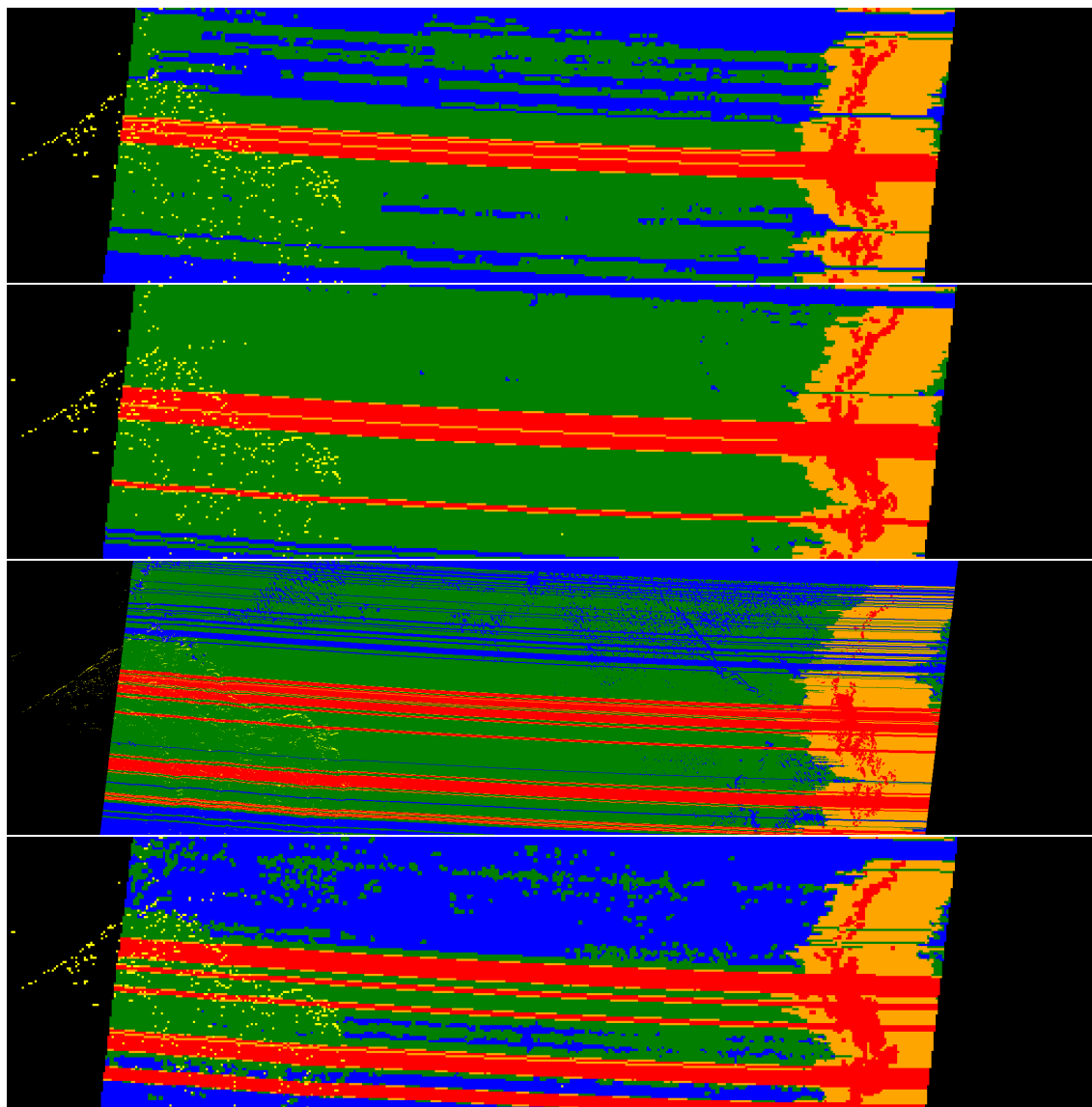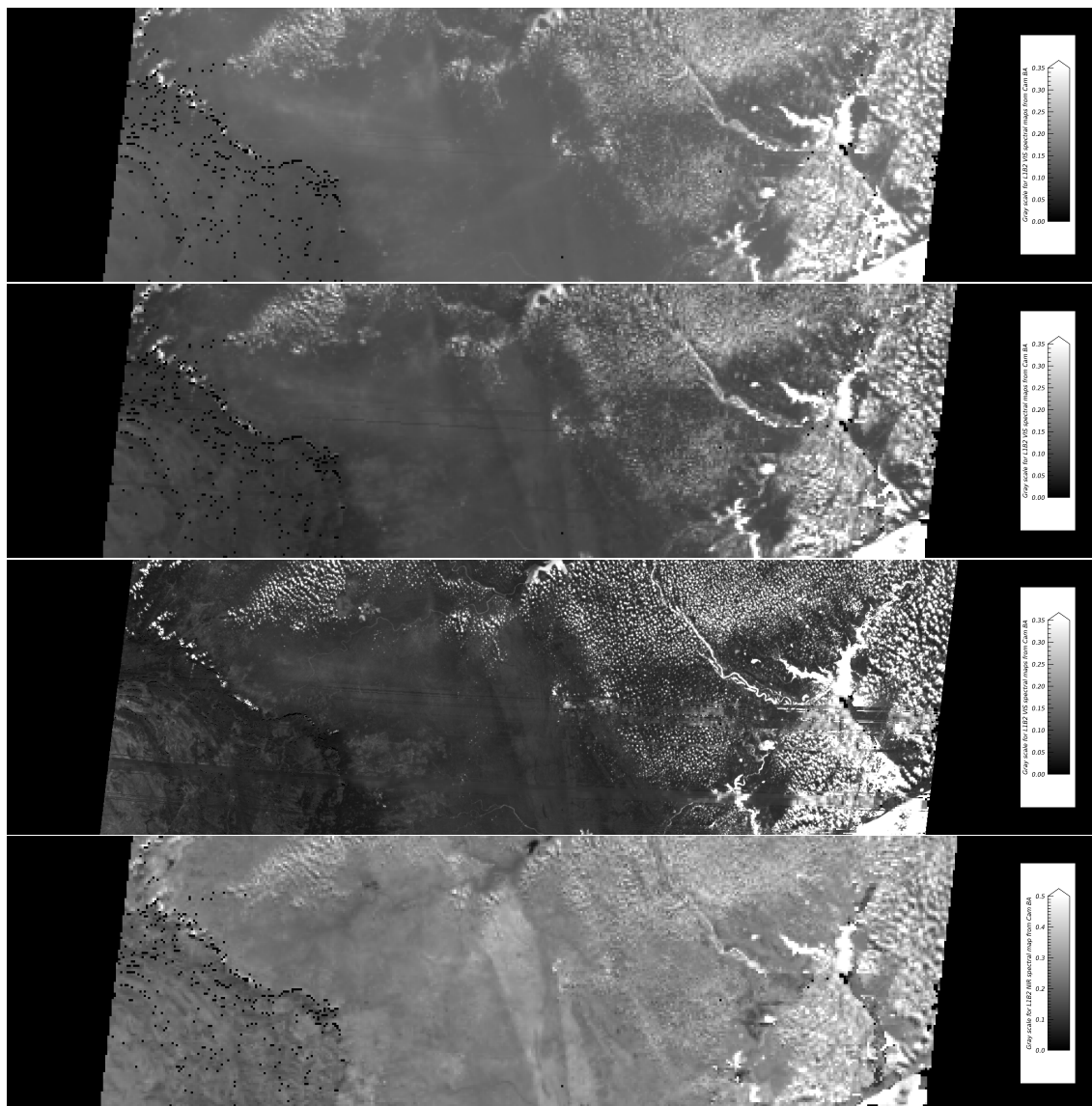

Figure 14. Maps of original RDQI values for GM, P168, O002111, B110 and Camera BA, in the Blue, Green, Red and NIR spectral bands, from top to bottom. Linear dimensions are the same as in Figure 1, and color coding is as described in the legend of Figure 10.

The original bidirectional reflectance factor product for these same 4 spectral bands are shown in Figure 15: it can be seen that missing values severely affect the usability of the data acquired by this camera.
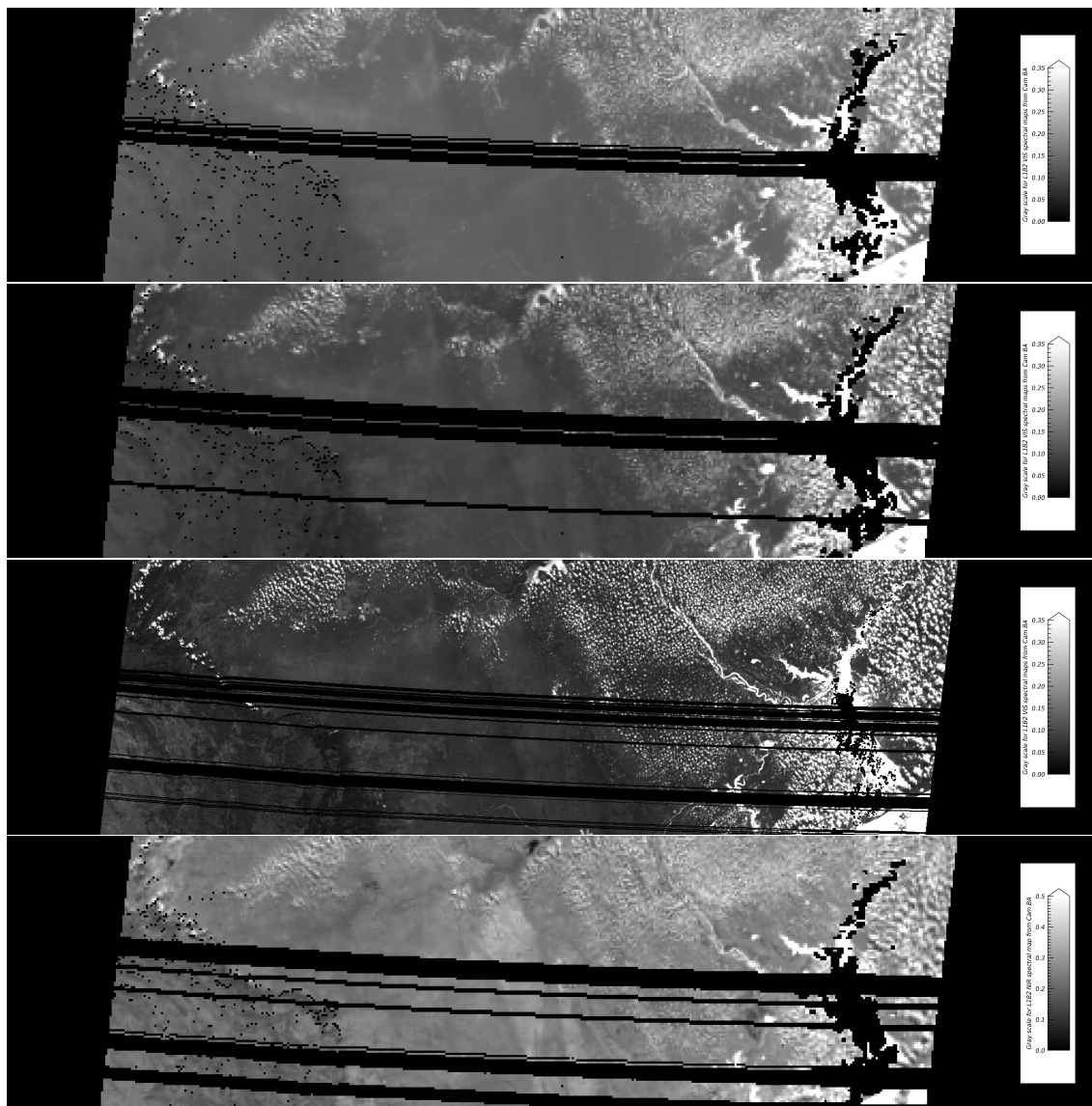
Figure 15. Maps of the original ToA bidirectional reflectance factor values for GM, P168, O002111, B110 and Camera BA, in the Blue, Green, Red and NIR spectral bands, from top to bottom. Linear dimensions are the same as in Figure 1, and gray scale stretching is the same as before: it assigns black to a null reflectance and white to reflectances of 0.35 or more in the blue, green and red spectral bands, or 0.50 in the near-infrared spectral band.

Figure 16 exhibits the updated bidirectional reflectance factor values in each of the 4 spectral channels after processing those
345   data as described above (exceptionally allowing for 8 successive attempts due to the significant overlap between the areas affected by missing values).

Figure 16. Maps of the updated ToA bidirectional reflectance factor values for GM, P168, O002111, B110 and Camera BA, in the Blue, Green, Red and NIR spectral bands, from top to bottom. Achieving this level of replacement required 8 iterations (see text for details). Linear dimensions are the same as in Figure 1, and gray scale stretching is the same as before: it assigns black to a null reflectance and white to reflectances of 0.35 or more in the blue, green and red spectral bands, or 0.50 in the near-infrared spectral band.

5   Algorithm evaluation

While the results exhibited in the previous section may be visually impressive, a key remaining question is to assess the numerical accuracy of this procedure. Since it is not possible to evaluate the pertinence of the replaced values when the original

350 data are missing, the process will be appraised by artificially introducing missing values in files that do not have missing lines in the first place, and then comparing the reconstructed values against the original ones.

The original valid data are first recorded separately, and missing values are inserted in MISR Global Mode data files for Path 168, Orbit 68050 and Block 110 between lines 30 and 34 of the Green spectral band in Camera CF, lines 100 to 110 of the Red spectral band in Camera AN, and lines 50 to 54 of the NIR spectral band in Camera DA. Figure 17 shows the corresponding

355 maps, where the black missing lines are clearly visible. Note the difference between those straight artificial missing and the curved lines that occur when data are missing following an instrument issue, as well as the rather larger areas affected.
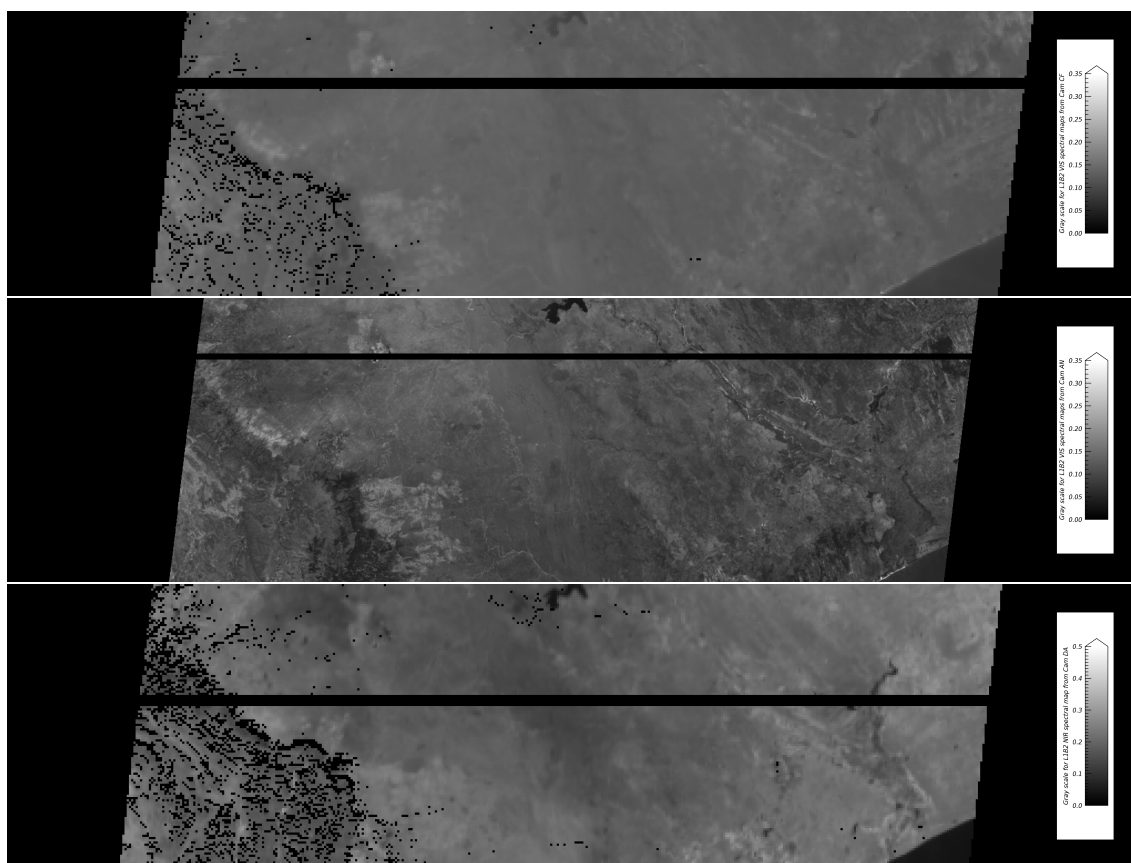
Figure 17. Maps of the MISR L1B2 data channels for the CF/Green, AN/Red and DA/NIR data channels of P168, O068050 and B110, where missing data have been inserted artificially for the purpose of testing the algorithm described in this paper. Linear dimensions are the same as in Figure 1, and gray scale stretching is the same as before: it assigns black to a null reflectance and white to reflectances of 0.35 or more in the blue, green and red spectral bands, or 0.50 in the near-infrared spectral band.

These data files are then processed as before, and the 3 left panels of Figure 18 exhibits the correlation statistics for the data channels that are best correlated with those that need to be updated. The 3 panels shown on the right compare the original values in abscissa to the reconstructed values in ordinates.
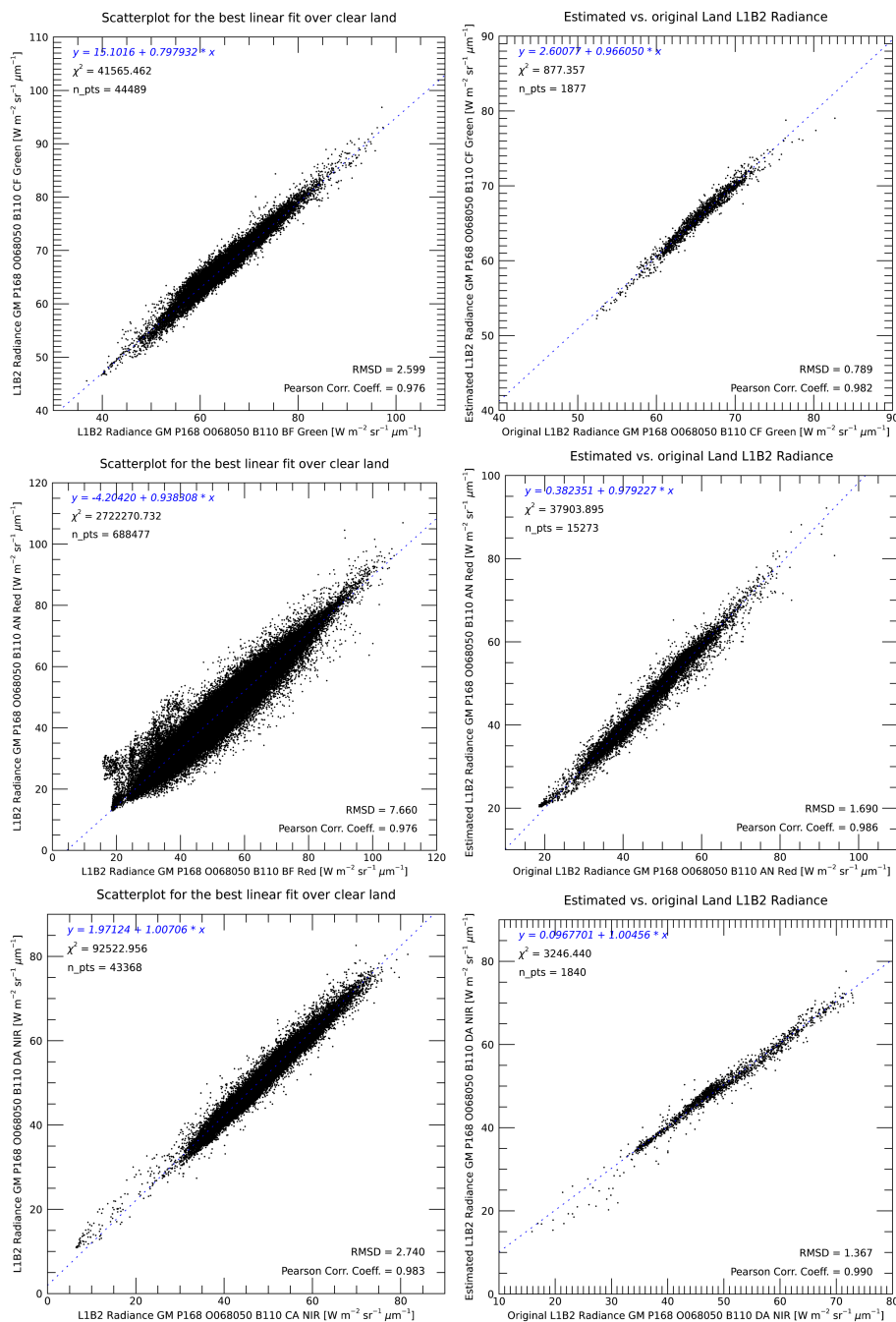
Figure 18. Scatterplots of the reconstructed versus the original values in the 3 data channels where missing values were artificially introduced as shown in 17: the top, middle and bottom panels concern the CF/Green, AN/Red and DA/NIR data channels, and in each case, the left panel shows the best predictor (abscissa) for the affected data channel (ordinate), while the right panel exhibits the relation between the original (abscissa) and the reconstructed (ordinate) values, in the areas where missing values were artificially introduced. Note that the ranges of values are different for each plot in order to optimize the view of the data.

360   Figure 19 exhibits the maps of those 3 data channels, where the missing data have been replaced by best estimates obtained as described above.
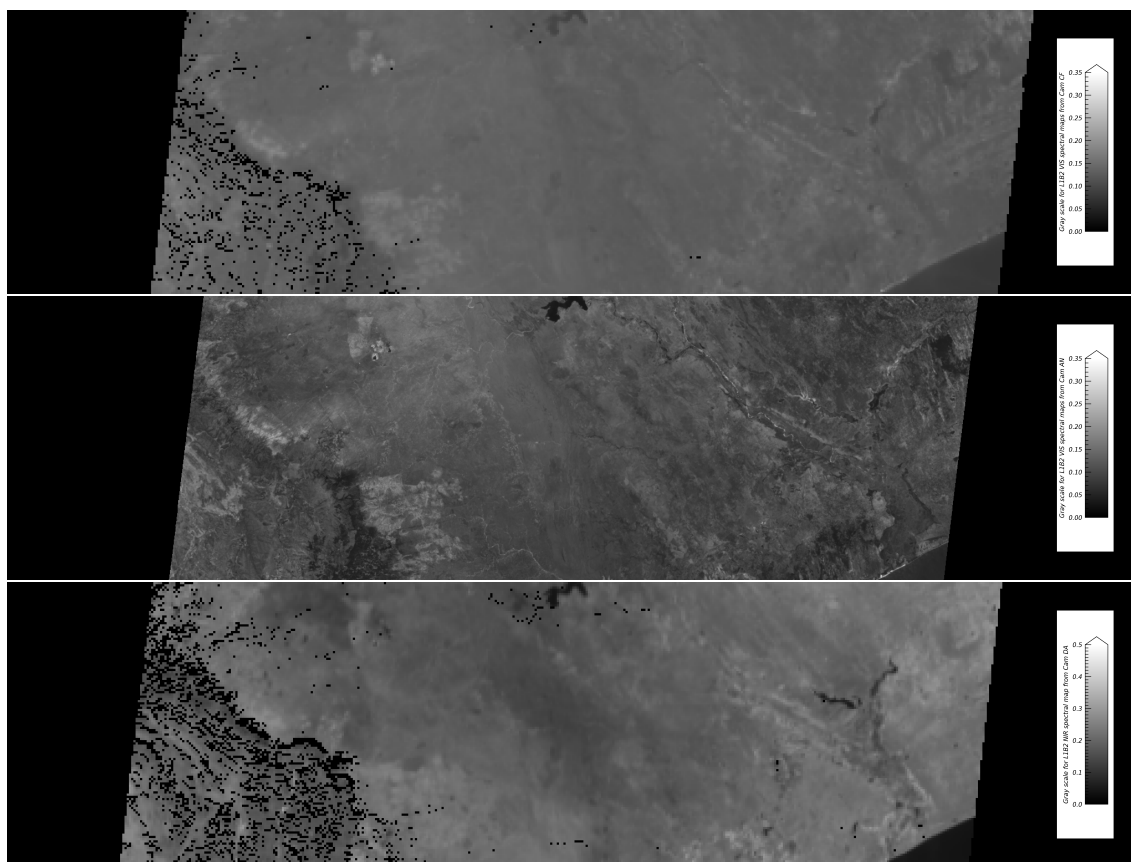
Figure 19. Maps of the MISR L1B2 data channels for the CF/Green, AN/Red and DA/NIR data channels of P168, O068050 and B110, where artificially inserted missing data have been replaced by best estimates, using the algorithm described in this paper. Linear dimensions are the same as in Figure 1, and gray scale stretching is the same as before: it assigns black to a null reflectance and white to reflectances of 0.35 or more in the blue, green and red spectral bands, or 0.50 in the near-infrared spectral band.

Similar results have been obtained in all other cases inspected. A second example is considered next, for the case of P168, O065487, B110 documented in Figures 10, 11, 12 and 13, which includes a significant cloud cover. In that case too, the original data channels CF/Green, AN/Red and DA/NIR do not suffer from any missing values, so the same patterns of missing data as
365   described above were artificially inserted in the original data, which were then processed as indicated above. Table 5 provides the number of missing points effectively replaced over clear land masses, the Root Mean Square Difference and the Pearson correlation coefficient between the original and the reconstructed data, as well as the χ^2, the unreduced chi-square statistic that expresses the goodness-of-fit of the linear relation between those to data sets. As noted before, the χ^2 statistics for the AN/Red data channel is higher due to the much larger number of pixel values involved. Note also that the number of clear land values

370 effectively replaced depends on the cloudiness present in the scene and therefore on the position of the artificially introduced lines of missing data with in the BLOCK.

Table 5. Performance statistics showing the high correlation between the original and the reconstructed MISR L1B2 data over clear land masses, when missing data are artificially introduced in the original data files, for Global Mode P168, O065487, B110. The numbers of clear land values involved depend on the cloudiness present in the scene, and therefore on the position of the artificially introduced lines of missing data with in the BLOCK.

Channel	# points	RMSD	PCC	χ^2
CF/Green	165	3.915	0.990	574
AN/Red	1136	2.415	0.990	2762
DA/NIR	600	2.632	0.930	3512

6 Conclusions

MISR's L1B2 Georectified Radiance Product (GRP) ellipsoid- or terrain-projected Global and Local Mode files occasionally include missing or poor data values, in one or more cameras, most often in multiple or all spectral bands. Their occurrence may

375 result from multiple causes, but the most frequent is the switching between Global and Local Mode. When a particular area is methodically acquired in Local Mode, as is the case for the Skukuza site in the Kruger National Park, South Africa, those BLOCKs are systematically contaminated by such missing data.

Every radiance measurement acquired by the MISR instrument is scaled into a 14-bit unsigned integer, to which a 2-bit Radiometric Data Quality Indicator (RDQI) is then appended. Missing data are assigned a specific 16-bit scaled radiance code

380 of `65523`, equivalent to a 14-bit code of `16380` with an RDQI of `3` attached, while measurements of dubious value (e.g., possibly affected by sun glint) are identified with an RDQI of `2`.

The process that generates these missing data points generally affects most or all of the 4 spectral bands but for only 1 or 2 (rarely 3) of the 9 cameras. However, due to the specific design of the instrument's focal planes, these missing lines affect different geographical locations in different spectral bands and cameras. Consequently, while a particular measurement may be

385 missing in a given camera and spectral band, measurements may be available in other spectral bands or cameras for the same location at essentially the same time. The high correlation between some of the data channels then permits the estimation of those missing values.

A set of software functions, written in the IDL[TM] language[1], was developed to implement the algorithms described above and process MISR L1B2 GRP data to replace those missing values. A selection of cases was exhibited to visually show that

390 the replacement process works well, and a series of tests, based on the artificial insertion of missing data (in files that do not contain any) permitted to show quantitatively the performance and accuracy of the procedure.

---

[1]IDL is a Trademark of Harris Geospatial Solutions, Inc.

As an aside, radiance measurements rated as 'poor', i.e., with an RDQI of 2, have often been found to cluster along those missing lines. These values could also be optionally be replaced by estimates, following the same procedure described above for missing data. The software allows this too, though in that case there is no obvious process for evaluating the results, i.e.,

395 there is no objective benchmark to decide whether the original value or the one suggested by the algorithm is closer to the nominal "true" value. The software does allow for the direct comparison between the original and the reconstructed values, through statistics and scatterplots. Users can experiment with both approaches and evaluate whether one or the other yields better or more appropriate results for their applications.

If missing or poor values are effectively replaced by estimates in the updated L1B2 data set, the RDQI assigned to the

400 corrected scaled radiances in L1B2 data buffers is set to 1.

The proposed method to replace missing and optionally poor data in MISR Global or Local Mode files is generally applicable, and has been implemented in the MISR-HR processing system. The materials in this paper are applicable to Version 2.1.x of the software, unless noted explicitly in the in-line documentation of the IDL functions. All MISR L1B2 input files are available from NASA's LaRC ASDC Distributed Active Archive Center (DAAC), and routines to explore and address those

405 issues are published on the first author's GitHub web page at `https://github.com/mmverstraete`.

Code availability. This paper describes Version 2.1.x of the IDL functions available in the following GitHub repositories, listed in order of increasing complexity:

- `https://github.com/mmverstraete/Macros;`
  `DOI: https://doi.org/10.5281/zenodo.3519843`

410
- `https://github.com/mmverstraete/Utilities;`
  `DOI: https://doi.org/10.5281/zenodo.3519846`

- `https://github.com/mmverstraete/MISR_Tools;`
  `DOI: https://doi.org/10.5281/zenodo.3519861`

- `https://github.com/mmverstraete/MISR_AGP;`
415
  `DOI: https://doi.org/10.5281/zenodo.3519897`

- `https://github.com/mmverstraete/MISR_RCCM;`
  `DOI: https://doi.org/10.5281/zenodo.3519901`

- `https://github.com/mmverstraete/MISR_L1B2;`
  `DOI: https://doi.org/10.5281/zenodo.3519989`

420 Each function contained in those repositories contains abundant in-line documentation to describe every step in the processing. An HTML file with properly formatted documentation for each function in a standardized setup is also available from the same repository. Functions in the more elaborate repositories may depend on routines defined in more basic repositories. It is thus recommended to download and install the contents of all repositories to have access to the full set of tools. Please remember also that those tools are under active development and may be updated in time.

425 *Data availability.* All MISR data, including the RCCM and the L1B2 radiance data products mentioned in this paper, were obtained from the NASA Langley Research Center's Atmospheric Science Data Center (ASDC). These data sets can be referenced as Diner et al. (1999b) (`https://doi.org/10.5067/Terra/MISR/MIRCCM_L2.004`) for RCCM and Jovanovic et al. (1999) (`https://doi.org/10.5067/Terra/MISR/MI1B2T_L1.003`) for L1B2, respectively. All MISR data products are openly available from the ASDC web site at: `https://eosweb.larc.nasa.gov/project/misr/misr_table`.

430 *Sample availability.* Samples of updated MISR L1B2 data products with missing data replaced by estimated values can be obtained by sending an email request to: `misrhrwits@gmail.com`.

Author contributions. Conceptualization, M.M.V.; Methodology, M.M.V.; Software, L.A.H. and M.M.V.; Validation, M.M.V., L.A.H., V.M.J.; Statistics on the prevalence of missing values: M.M.V.; Original and final draft, M.M.V.; Review and editing, L.A.H., V.M.J.; Supervision, M.M.V.

435 *Competing interests.* The authors declare no conflict of interest and the indirect funding sponsor (NRF) had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Disclaimer. The IDL source code software mentioned above (obtainable from the GitHub web site) is made available under the MIT license, which states, in part, that "The software is provided *as is*, without warranty of any kind, express or implied, including but not limited to the warranties of merchantability, fitness for a particular purpose and noninfringement. In no event shall the authors or copyright holders be
440 liable for any claim, damages or other liability, whether in an action of contract, tort or otherwise, arising from, out of or in connection with the software or the use or other dealings in the software." Please read carefully and abide by the full license before downloading and using this software.

References

450 Bull, M., Matthews, J., McDonald, D., Menzies, A., Moroney, C., Mueller, K., Paradise, S., and Smyth, M.: Data Products Specifications, Tech. Rep. JPL D-13963, Revision S, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, https://eosweb.larc.nasa.gov/project/misr/dps, 2011.

Diner, D. J., Bruegge, C. J., Martonchik, J. V., Ackerman, T. P., Davies, R., Gerstl, S. A. W., Gordon, H. R., Sellers, P. J., Clark, J., Daniels, J. A., Danielson, E. D., Duval, V. G., Klaasen, K. P., Lilienthal, G. W., Nakamoto, D. I., Pagano, R. J., and Reilly, T. H.: MISR: A

455 Multiangle Imaging SpectroRadiometer for Geophysical and Climatological Research from EOS, IEEE Transactions on Geoscience and Remote Sensing, 27, 200–214, https://doi.org/10.1109/36.20299, 1989.

Diner, D. J., Beckert, J. C., Reilly, T. H., Ackerman, T. P., Bruegge, C. J., Conel, J. E., Davies, R., Gerstl, S. A. W., Gordon, H. R., Kahn, R. A., Martonchik, J. V., Muller, J.-P., Myneni, R. B., Pinty, B., Sellers, P. J., and Verstraete, M. M.: Multi-angle Imaging SpectroRadiometer (MISR) Instrument Description and Experiment Overview, IEEE Transactions on Geoscience and Remote Sensing, 36, 1072–1087,

460 https://doi.org/10.1109/36.700992, 1998.

Diner, D. J., Asner, G. P., Davies, R., Knyazikhin, Y., Muller, J.-P., Nolin, A. W., Pinty, B., Schaaf, C. B., and Stroeve, J.: New Directions in Earth Observing: Scientific Applications of Multiangle Remote Sensing, Bulletin of the American Meteorological Society, 80, 2209–2228, https://doi.org/10.1175/1520-0477(1999)080<2209:NDIEOS>2.0.CO;2, 1999a.

Diner, D. J., Di Girolamo, L., and Clothiaux, E. E.: Level 1 Cloud Detection Algorithm Theoretical Basis, Tech.

465 Rep. JPL D-13397, Revision B, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA, https://doi.org/10.5067/Terra/MISR/MIRCCM_L2.004, https://eospso.gsfc.nasa.gov/atbd-category/45, 1999b.

Diner, D. J., Verstraete, M. M., and Martonchik, J. V.: Foreword to Special Section on MISR, IEEE Transactions on Geoscience and Remote Sensing, 40, 1447–1448, https://doi.org/10.1109/TGRS.2002.802777, 2002.

Diner, D. J., Braswell, B. H., Davies, R., Gobron, N., Hu, J., Jin, Y., Kahn, R. A., Knyazikhin, Y., Loeb, N., Muller, J.-P., Nolin, A. W., Pinty,

470 B., Schaaf, C. B., Seiz, G., and Stroeve, J.: The value of multiangle measurements for retrieving structurally and radiatively consistent properties of clouds, aerosols, and surfaces, Remote Sensing of Environment, 97, 495–518, https://doi.org/10.1016/j.rse.2005.06.006, 2005.

Jovanovic, V. M., Lewicki, S. A., Smyth, M. M., Zong, J., and Korechoff, R. P.: Level 1 Georectification and Registration Algorithm Theoretical Basis, Tech. Rep. JPL D-11532, Revision D, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA,

475 USA, https://doi.org/10.5067/Terra/MISR/MI1B2T_L1.003, https://eospso.gsfc.nasa.gov/atbd-category/45, 1999.

Liu, Z., Verstraete, M. M., and de Jager, G.: Handling outliers in model inversion studies: a remote sensing case study using MISR-HR data in South Africa, South African Geographical Journal, 100, 122–139, https://doi.org/10.1080/03736245.2017.1339629, 2018.

Mahlangu, P., Mathieu, R., Wessels, K., Naidoo, L., Verstraete, M. M., Asner, G., and Main, R.: Indirect Estimation of Structural Parameters in South African Forests Using MISR-HR and LiDAR Remote Sensing Data, Remote Sensing, 10, 1537–1567,

480 https://doi.org/10.3390/rs10101537, 2018.

Van den Hoof, C., Verstraete, M. M., and Scholes, R. J.: Differing Responses to Rainfall Suggest More Than One Functional Type of Grassland in South Africa, Remote Sensing, 10, 2055–2077, https://doi.org/10.3390/rs10122055, 2018.

Verstraete, M. M., Hunt, L. A., Scholes, R. J., Clerici, M., Pinty, B., and Nelson, D. L.: Generating 275-m Resolution Land Surface Products From the Multi-Angle Imaging SpectroRadiometer Data, IEEE Transactions on Geoscience and Remote Sensing, 50, 3980–3990,

485 https://doi.org/10.1109/TGRS.2012.2189575, 2012.

Verstraete, M. M., Hunt, L. A., De Lemos, H., and Di Girolamo, L.: Replacing Missing Values in the Standard MISR Radiometric Camera-by-Camera Cloud Mask (RCCM) Data Product, Earth System Science Data Discussion, xx, yy–yy, https://doi.org/https://doi.org/10.5194/essd-2019-77, 2019.

Yang, Y., Di Girolamo, L., and Mazzoni, D.: Selection of the automated thresholding algorithm for the Multi-angle Imaging SpectroRadiometer Radiometric Camera-by-Camera Cloud Mask over land, Remote Sensing of Environment, 107, 159–171, https://doi.org/10.1016/j.rse.2006.05.020, 2007.

Zhao, G. and Di Girolamo, L.: A cloud fraction versus view angle technique for automatic in-scene evaluation of the MISR cloud mask, Journal of Applied Meteorology, 43, 860–869, https://doi.org/10.1175/1520-0450(2004)043<0860:ACFVVA>2.0.CO;2, 2004.