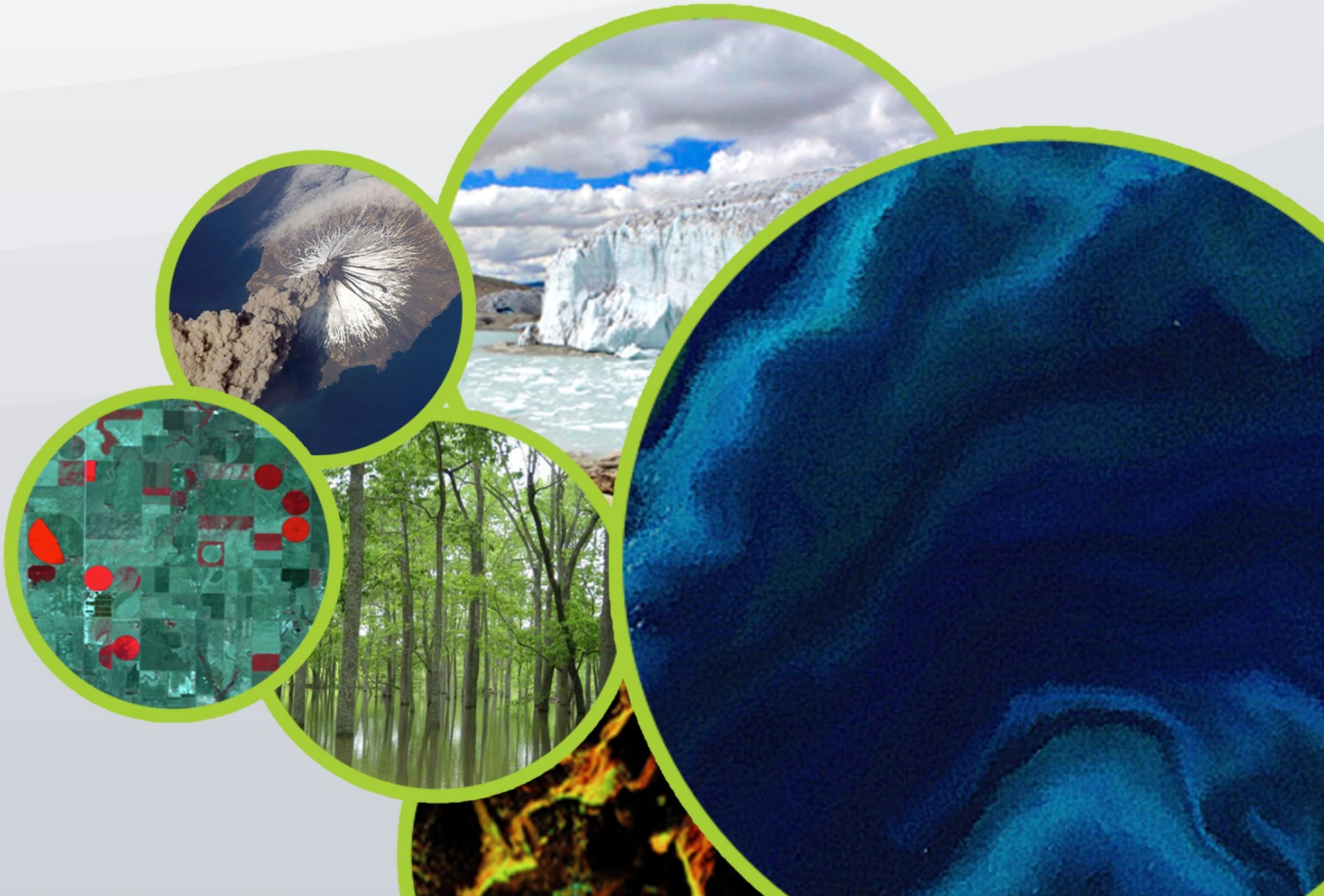
CEOS Committee on Earth Observation Satellites



Moderate Resolution Sensor Interoperability Initiative

Version 1.0, October 2017

Moderate Resolution Sensor Interoperability Initiative

Version 1.0, October 2017

Introduction

The Moderate Resolution Sensor Interoperability (MRI) initiative addresses the CEOS strategic objective to support complementarity and compatibility among moderate resolution (10-100 meter) Earth Observation (EO) sensors and data received from them. It also identifies data production and use issues that must be considered for the successful implementation of multi-sensor, interoperable time series. To this end, the MRI team has developed a Framework that can be used to document parameters and systematically address issues with EO data interoperability, making optimal use of the increasing number of data streams available in the moderate resolution class. Although an initial case study of the use of the Framework has focused on Landsat/Sentinel-2 interoperability, the Framework is specifically designed to be applied to determining interoperability between any moderate resolution sensors or datasets.

The overarching goal of interoperability is to enable the use of complementary sensors to achieve a single, coherent data stream that characterises change on the Earth's surface through time. As EO data products are developed—including high-level products that span long time periods--the integration of these products requires verification and validation of their interoperability in a way that addresses fundamental questions, such as: *When can—or can't—these products be compared, and under what conditions?*

Many factors, including geolocation accuracies, spatial resolution, radiometric and spectral inconsistencies, atmospheric effects, natural variability of surface features, and instrument differences, can cause variability in data through time. We have worked to identify those sources of variability that scientists and land managers should consider when comparing multi-sensor products. Additionally, we emphasise that interoperable metadata are critical for data discovery and access, and for the maintenance and updating of image databases such as dense, multi-sensor time series in data cubes.

Recognizing the importance of defining a Framework for moderate resolution interoperability, the MRI initiative originated as one of the two 2017 CEOS Chair Initiatives, and was established as an activity of the CEOS Land Surface Imaging Virtual Constellation (LSI-VC). The MRI team relied heavily on CEOS members and activities associated with LSI-VC, including the Working Group on Calibration and Validation (WGCV), Working Group on Information Systems & Services (WGISS), Global Forest Observations Initiative (GFOI), Group on Earth Observations Global Agricultural Monitoring Initiative (GEOGLAM), Future Data Access & Analysis Architectures (FDA), and CEOS Analysis Ready Data for Land (CARD4L) [1]. These represented both space agencies and user communities. As examples, the FDA [2]and CARD4L efforts complement the objectives of this initiative from the perspectives of space agencies, whereas GEOGLAM and GFOI activities are in a strong position to provide user feedback on acceptable application-specific requirements.

For 2017, the MRI initiative undertook an initial case study, summarised in a later section, which involves documenting the application of the Framework to an example of Landsat/Sentinel-2 data interoperability. The intent of this and future case studies (2018 onward) implemented under the MRI initiative is to document not only the application of the Framework in assessing interoperability, but also to publish and communicate to the user community the objectives and intended uses of the interoperable products, former and current activities that support product development, and the status of these activities.

The MRI Framework: A Conceptual Evaluation Tool

There are three underlying requirements for achieving multi-sensor interoperability:

- **Provide methodologies to determine the uncertainties for input and combined products**. The goal is to understand the uncertainties at each step of production; for example, by allocating uncertainty to at-sensor products, atmospheric corrections, illumination and view angle corrections, band difference corrections, and classification, as appropriate.
- **Determine the acceptable uncertainties for specific applications**, as identified by the user community.
- **Establish interoperable per-scene and per-pixel metadata for use in data discovery and as analytical filters**. Metadata can be used to identify available scenes, data quality, and data gaps where clear pixels are not available.

The continued application of the MRI Framework to multiple sensors or datasets in order to assess their interoperability should result in:

1. **Recommendations to data producers for product evolution to meet interoperability requirements;**
2. **Good practices and guidance for the user community in adapting and using multiple sensors products within single data streams.**

The Framework is a first-step interoperability analysis tool that will evolve as products evolve and as case studies are implemented. We seek feedback from both data producers and data users to establish and expand on these concepts, alternates, thresholds and targets. The producer and user communities are in a period of rapid change as dense time series of 10-100m, analysis-ready data become increasingly easy to access.

The scope of the initiative is restricted to moderate resolution sensors designed for global monitoring of land and associated water features. Appendix A provides a high-level summary of many CEOS agency 10-100m data sources [3-18]. We encourage CEOS agencies to explore other sensor interoperability combinations, including other resolutions, and non-land related applications using the Framework. While the initial focus is on optical sensors with an emphasis on surface reflectance products, the Framework will accommodate inclusion of Synthetic Aperture Radar (SAR) and higher level products.

The MRI Framework brings an ordered, systematic approach to determining the degree to which datasets or products from different EO sensors, are—or have the potential to be—interoperable. It aids in identifying data and metadata characteristics for data products that may affect multi-sensor applications.

As seen in Table 1, the Framework is organised into four components: general metadata, per-pixel metadata, data measurements, and geolocation. Within each of these components are a set of interoperability concepts (items) for which alternative solutions can be identified, threshold verifications discussed, and target next steps proposed. As reflected in the Framework, threshold refers to the agreed-upon minimum value or characteristic that is acceptable for a specific application, whereas the target represents the goal (ideal characteristic or measure) for which producers and users are striving. Comparison of actual measurements to predetermined thresholds/targets is documented for the items of each component as part of the verification process, followed by a listing of the next steps required to address any shortfalls.

Table 1. The MRI Framework

Component	Items	Threshold	Target	Verification	Next Steps
General Metadata	Coordinate reference system Reference grid accuracy Geometric accuracy Spectral bands Spectral response curves Radiometric accuracy Revisit time & lifetime Field of view Mean local time				
Per-pixel Metadata	Cloud cover Cloud shadow Land/water masks Snow/ice masks DEM Terrain shadow mask Illumination and viewing geometry Data quality				
Data Measurements	Measurements Measurement normalisation Aerosol/water vapor/ozone corrections Spectral band difference corrections				
Geolocation	Geometric corrections Resampling				

The use of the MRI Framework assumes that interoperability of two products is 'application specific.' The same two products (e.g., one Landsat 8 and one Sentinel-2) may meet interoperability specs for class changes such as forest to non-forest, but not meet requirements for detecting forest degradation. Each completed MRI Framework needs to clearly identify the application it serves.

Thresholds and the corresponding targets for the sensors being assessed will be based on those application-specific requirements determined by the participating producers and users. Following verification, next steps are determined and will very likely be unique to that Framework for that application.

A major strength of the Framework is that it captures the process of evaluating these parameters. The record of the development of the resulting Framework provides a significant amount of information to the community at large. Agreement on the characteristics (thresholds) of the complete range of parameters of any two sensors in the context of a specific application not only increases understanding between those specific producers and users, but can be used as a model for similar Frameworks developed subsequently by other members of the community. More importantly, the community will be encouraged to continue to populate and refine each Framework.

We anticipate that the MRI Framework will evolve as products, technology, and user applications evolve. Some sensors are very similar and more easily made interoperable. Other sensors have significant identifiable differences for which cross-calibration factors must be produced to enable the datasets to be used together. Interoperability solutions may include one or both of two common paths:

1. Changes to operational products or post-processing methodologies to create interoperable products;

2. Accommodation to inherent differences between products.

Changes to products include radiometric cross-calibration to standard references, and acceptance of compatible geographic reference grids, digital elevations models (DEMs), bidirectional reflectance distribution function (BRDF) models, and atmospheric models. Where appropriate, the same models and references can be adopted; otherwise, differences between the references and methodologies need to be quantified and documented. These changes are considered **harmonization**.

Accommodations to inherent differences include pixel size, field of view, different spectral response curves, view and solar angle variations, and available bands. Accommodation may include resampling and scaling products to provide a comparable merged product, or robust and flexible application methodologies designed to accommodate the differences. Many of the new trending methodologies, such as Continuous Change Detection and Classification (CCDC), take advantage of the increased temporal density to implement robust outlier logic [19-27]. Accommodation has the best results if all possible attempts are made to harmonize the products as early as possible in the product production flow, to take advantage of space agency expertise and to focus user expertise requirements on applications rather than data preprocessing. Modification of the products to create a consistent time series product is considered **homogenization**.

The CARD4L product family specifications strive to create products that meet the criteria for single-sensor, intra-product analysis to minimise the specialised knowledge needed to preprocess the data prior to multi-sensor analysis. If two products are sufficiently similar—for example, possessing similar pixel size and spectral bands—then they can be used interoperably for many applications. Otherwise, steps are needed to harmonize and homogenize the products to make them interoperable. These homogenization steps may be application specific. Both harmonization and homogenization are independent of implementation. The implementation may include the creation of discrete products or may be virtual implementation within a model. Key to success is knowledge of the uncertainty of the individual source products, combined data streams, and user requirements.

The results of the MRI initiative will support the CARD4L description document [1]. CARD4L identifies information needed for a product to be considered Analysis Ready Data (ARD) [28] for the user community. The CARD4L Product Family Specifications define threshold and target information needed to provide a minimal ARD product [29-31]. Thresholds identify minimum current and verifiable product requirements. Targets identify a product evolutionary path—how agencies can continue to improve products.

The threshold information is either now in use in operational products or is known and documented within the community. Products that meet all CARD4L threshold requirements are considered analysis-ready for scientific analysis or decision making.

Products that meet CARD4L target requirements further reduce overall product uncertainties, increase accuracy, and enhance broad-scale applications requiring interoperable products. The target information is often under development and not widely implemented in operational products. The target requirements anticipate continuous improvement of methods and evolution of community expectations, which are both normal and inevitable in a developing field. Over time, *target* specifications may (subject to due process) become accepted as *threshold* requirements.

Table 2 summarises consequences associated with moderate resolution sensor interoperability items for each of the four main components in the MRI Framework. These will be discussed in further detail in the component sections that follow, as will current threshold requirements and future targets for improvement, as well as initial recommendations.

Table 2. A Summary of MRI Framework Items with Interoperability Consequences

Component	Items	Consequences
General Metadata	Coordinate reference system	Different pixel sizes, origins, and projections require the data be spatially resampled. Larger differences will cause larger changes in radiometry.
	Reference grid accuracy	Different and less accurate reference grids will cause greater uncertainty in spatial alignment of pixels.
	Geometric accuracy and temporal consistency	Absolute spatial root-mean-square error (RMSE) distributions and temporal spatial consistency for each sensor are needed to establish the joint distribution for multi-sensor time series.
	Spectral bands	Different bands between sensors require methodologies that can adapt to band availability.
	Spectral response curves	Different spectral response curves may exist for similar bands between sensors, causing land surface dependent variability.
	Radiometric accuracy	Biases and uncertainty need to be minimised between sensors using spectrally uniform references and opportunities for simultaneous imaging.
	Revisit time & lifetime	Multi-sensor datasets can increase the length of the time series and increase the density of the time series.
	Field of view	The greater the swath width or field of view, the more importance needs to be placed on illumination and viewing geometry and on the DEM used.
	Mean local time	Different mean local times between sensors and over the lifetime of a sensor will be reflected in the reflectance values.
Per-pixel Metadata	Cloud cover	Verified and validated cloud masks are needed to estimate radiometric contamination for specific and known cloud characteristics.
	Cloud shadow	Verified and validated cloud shadow masks are needed to estimate radiometric contamination for specific and known cloud characteristics.
	Land/water masks	Land and water masks provide useful information for other radiometric corrections.
	Snow/ice masks	Snow/ice masks assist in pixel filters and illumination and viewing geometry corrections. Known confusion, such as with clouds, needs to be documented.
	DEM	The required accuracy of the DEM is dependent upon the corrections implemented, swath width, and pixel size.
	Terrain shadow mask	Terrain shadow masks are needed to estimate radiometric contamination associated with shadows. Known confusion, such as with water and cloud shadow, needs to be quantified.
	Illumination and viewing geometry	Solar illumination angles are needed for reflectance calculations. Solar and view angles are needed for BRDF-related corrections.
	Data quality	No data, saturated, contaminated, terrain occlusion pixels need to be identified.

Component	Items	Consequences
Data Measurements	Measurements	Absolute calibrated measurement units with or without corrections below.
	Measurement normalisation	Radiometry viewed through time is significantly impacted by variation in solar and viewing angles.
	Aerosol/water vapor/ozone corrections	Different atmospheric models can introduce significant between-sensor variability.
	Spectral band difference corrections	Different spectral response curves will introduce differences between products.
Geolocation	Geometric corrections	Residual misregistration between images introduces variability in the radiometry measurements.
	Resampling	The number and type of spatial resampling will impact the radiometric signal.

The MRI Framework: Item Analyses

The four components in the MRI Framework—general metadata, per-pixel metadata, data measurements, and geolocation—are discussed in more detail below, as an item-by-item analysis. Concepts and alternatives are explored as well as specific thresholds and next steps toward meeting targets to further improve multi-sensor interoperability. General and per-pixel metadata typically serve as input to radiometric and geometric methodologies. Recommendations and notes (*in italics*) are given at the end of each item discussion.

General Metadata

General metadata are provided at the product or scene level. Product-level metadata should, at a minimum, be documented in user guides and published literature. Scene-level metadata should be available in a machine-readable format, since this information is needed to filter specific scenes for inclusion in analysis. Table 3 identifies MRI Framework thresholds and targets for general metadata.

The overall threshold objective for general metadata is to bundle machine-readable metadata with the product, while the target objective is the adoption of either the Open Geospatial Consortium (OGC) metadata standard or the ISO 19115-2 standard "Geographic information – Metadata – Part 2: Extensions for imagery and gridded data." The threshold objective of machine-readable metadata is met by most products.

Table 3. General Metadata

Component	Items	Threshold	Target
General Metadata	Coordinate reference system	Document pixel sizes, origins, and map projections in machine readable format.	Document in standardised metadata format. When practical, establish common origins and map projections.

	Reference grid accuracy	Document absolute accuracy of reference data. Reference grid uncertainty contribution to geometric accuracy should be minimised.	Document relationships among reference databases in operational use. Share reference databases when possible. Adopt common accuracy metric.
	Geometric accuracy	Document uncertainty of each individual product and the methodologies used.	Document in standardised metadata format. Total uncertainty when combined with reference grid uncertainty should be on the order of 1/3 pixel. Adopt common accuracy metric.
	Spectral bands	Document available bands in machine-readable metadata.	Document in standardised metadata format. Quantify benefits provided by additional bands.
	Spectral response curves	Document spectral response curves in public literature.	Document spectral response curves in standardised metadata and in CEOS Mission, Instruments, and Measurements (MIM) database.
	Radiometric accuracy	Document biases and uncertainty in public literature. Acceptable accuracy is application specific.	Document total error budget and temporal consistency for product families in metadata. Continue to improve product accuracy and to understand application requirements.
	Revisit time & lifetime	Document revisit time and active lifetime in public literature. Interoperability goal is to achieve 7-day cloud-free revisit time.	Identify critical time periods and regions. Encourage access to historical archives. Extend interoperable time series globally to the beginning of the Landsat MSS period (1972) or earlier.
	Field of view	Document field of view. High-level products need to account for different viewing geometries.	Quantify radiometric uncertainty associated with off-nadir viewing angles.
	Mean local time	Document mean solar time. High level products need to account for different solar geometries.	Quantify uncertainty associated with different solar geometries between missions and through the life of the mission.

Coordinate Reference System

Coordinate reference systems are defined for product families. This information should be stored with products and in the CEOS MIM database [4]. Convergence by data production agencies toward a common nested tile system would reduce the need to resample data. It is understood that national requirements may dictate map projections and that cost, performance, and latency requirements limit on-demand selection of map projection. The OGC Discrete Global Grid Systems (DGGS) should be investigated as an alternative for common reference grids [32].

> *Users creating interoperable products need to reproject to either the smaller or larger pixel size, depending upon application requirements and common origin and projection, either through creating a resampled copy of the data or using an 'on-the-fly' approach. The consequences of reprojection/resampling need to be understood.*

Reference Grid Accuracy

The absolute geometric accuracy can be no better than the accuracy of the reference database [33-37]. If different reference databases are used, the uncertainty of the reference databases must be added to the geometric accuracy to estimate the geometric accuracy of the multi-sensor dataset. Data producers need to coordinate to minimise differences between reference grids.

For example, the reference grid accuracy for Global Land Survey (GLS) reference used for all Landsat 1-8 products is 17m CE95; Global Reference Image (GRI), which is under construction in 2017, is 9.5 CE95 [37]. The quality of the GLS reference grid is highly variable, with maximum offsets of over 36m.

> *Date producers need to minimise the absolute error of individual reference grids. Whenever possible, reference grids should be shared across sensors.*

Geometric Accuracy

Documentation of the geometric accuracy of each image permits selection of the images to meet an accuracy threshold for specific applications [38-40]. Estimates of geometric accuracy permit the identification of images that may not be stackable without further registration. Estimates of per image geometric accuracy are only available for images that are precision registered to a reference grid. Otherwise, accuracy estimates are based on systematic models and are not tied to ground references. Newer sensors have much-improved ephemeral location data acquired with the images that permit accurate systematic models comparable to precision models.

For example, the geometric accuracy for Landsat Operational Land Imager (OLI) is 14m CE95 in relation to the GLS reference grid. This accuracy estimate is relevant for image-to-image stacking of Landsat data using the GLS reference grid. For Sentinel-2, the relative accuracy is 10 m 2σ. Predicted Landsat 8 to Sentinel-2 uncertainty is 26m 2σ. Once the Sentinel-2 GRI is available and Landsat data are reprocessed to use GRI, a registration accuracy on the order of 10m 2σ is anticipated.

> *Data producers need to reduce RMSE to the theoretical minimum per sensor in relation to reference grid.*

Spectral Bands

Sensors have different available bands. In some multi-sensor scenarios, unique bands such as the aerosol, cirrus, red edge and thermal bands, will exist in addition to the core common bands such as visible and near-infrared (VNIR), near-infrared (NIR) and short-wave infrared (SWIR) bands. Merged datasets may or may not include unique bands for some sensors, depending on individual application requirements. Interoperability should never justify the exclusion of information that can improve an application result. However, the application methodology needs to accommodate the differences.

> *The bands available for analysis will differ among sensors. Users creating interoperable products should, when possible, provide the user community with the richest set of alternatives, even when it creates a discontinuity in the time-series record for some bands.*

Spectral Response Curves

Specific sensor bands, albeit nominally similar, will have unique spectral response curves. The differences between spectral responses may cause significant variability within time series. These differences may be exacerbated in derived products that use ratios such as the Normalized Difference Vegetation Index (NDVI) [41, 42].

> *Spectral response curves should be stored in the CEOS MIM database [4]. Spectral Band Adjustment Factors (SBAF), discussed below in the radiometry subsection, can be used with caution to accommodate spectral response differences.*

Radiometric Accuracy

The within-sensor calibration and multi-sensor cross-calibration of at-sensor data using geometric, instrumented, and pseudo-invariant calibration sites is fundamental to the production of radiometrically and geometrically accurate products. At-sensor uncertainty is a known component of the total error budget of higher level products. The relative accuracy is critical for interoperability. For example, the radiometric differences between Landsat 8 and Sentinel-2 are approximately 2% [43]. Understanding the contribution of atmospheric and bidirectional reflectance corrections to product uncertainty is important to establishing the total uncertainty as high-level products, such as surface reflectance, are more frequently used within user applications.

> *Data producers should coordinate cross-calibration of at-sensor products using uniform reference surfaces to minimise overall biases between sensors.*

Revisit Time & Lifetime

The rationale for multi-sensor interoperability is to extend the length and density of time-series datasets. Longer time series permit the establishment of earlier baselines and the identification of periodic changes. The harnessing of as much data as possible from multiple complementary sensors enables a richer set of data to be used, and additionally to fill in gaps left by cloud and other data issues, allowing denser time series [44]. Denser time series allows the study of changes that may be related to phenology, erosion, and other Earth processes, fostering a better understanding of natural and anthropogenic impacts.

The creation of multi-sensor time series extends time series and increases the temporal density at the cost of increased variability in the resulting time series.

> *Multi-sensor time series can consist of streams of measurement data, such as reflectance, or intervals of higher level products, such as classed data. Users can select sensors and products as appropriate to meet their analysis goals. Examples include the interoperable use of classed data derived from SAR and optical sensors.*

Field of View

A wider swath width reduces the revisit time if other orbital parameters are the same. However, the greater the swath width, the more importance needs to be placed on illumination and viewing geometry and on the DEM used for corrections. Pixel size is a function of field of view. Off-nadir pixels are inherently larger than nadir pixels. This difference is accommodated in product production, but can result in radiometry variability.

Field of view corrections need to also accommodate pointable sensors, such as ASTER and SPOT, which acquire operational data while pointing off-nadir.

> *The requirement of solar angle, viewing angle, atmospheric corrections, and orthorectification with high-resolution DEMs is more critical for larger swath widths with larger off-nadir views, particularly with multiple sensors that have different nadir lines.*

Mean Local Time

Mean solar time is a component of the viewing and solar angle corrections. Mean local time for polar-orbiting, sun-synchronous missions varies from 09:30 to 10:30. Differences in mean local time contribute significantly to data measurement variability. During the 27-year lifetime of the Landsat 5 mission, the mean local time varied by over an hour, which caused changes in the solar angle of over 10° [45].

> *View and solar angle corrections need to accommodate any changes in mean local time.*

Per-pixel Metadata

The inclusion of per-pixel metadata within data products has increased importance with the distribution of higher level, analysis-ready data. Cloud, shadow, quality, and other information available at the pixel level can be used directly within application models. However, even though much of the per-pixel metadata currently available is not validated, they can provide important information if used with caution. Some per-pixel metadata are derivative products produced from and distributed with the data product; others are external products, such as DEMs and atmospheric model inputs, used in the processing/derivation of the data product. External products have their own uncertainties that must be understood.

General, scene-level, metadata aggregations of per-pixel metadata, such as cloud cover, data quality, snow/ice, and solar angle, are often available in searchable databases. These metadata can be used to reject scenes from inclusion in time series that do not meet application criteria.

Interoperability requirements include documented algorithms and accuracies for cloud cover and shadow, land, water and vegetation, terrain shadow, atmospheric model inputs including aerosols, saturation, and other data qualities. Each of these per-pixel metadata sets is a function of a model that has its own confidence estimates, which will contribute toward the error budget of products. The level of validation and verification of these metadata is highly variable and needs to be documented. Differences in these per-pixel metadata vary between products and can add noise to the multi-sensor time series.

Currently, there is little standardization of per-pixel metadata by content, structure, or implementation within databases, including Data Cube methodologies. Interoperable per-pixel metadata provide an opportunity for quickly identifying where clear pixels are not sufficiently available and where more optical or SAR data are needed. The adoption or sharing of a common methodology that can be applied to different sensors would provide consistent metadata for multi-sensor data streams. Table 4 summarises Framework thresholds and targets for items associated with the per-pixel metadata component.

Table 4. Per-pixel Metadata

Component	Items	Threshold	Target
Per-pixel Metadata	Cloud cover	Document cloud definition and methodology, including treatment of cirrus clouds and cloud edges. Document potential confusion with other classes, such as sand, snow, and ice.	Verify and validate cloud masks. Include opacity and probability estimates. Investigate new bands needed to optimise estimates. Quantify confusion with other classes. Adopt common methodology and standards for use on multiple sensors.
	Cloud shadow	Document cloud shadow methodologies. Document potential confusion with other dark objects such as water and terrain shadow.	Verify and validate cloud shadow masks. Quantify confusion with other dark objects. Adopt common methodology and standards for use on multiple sensors.
	Land/water masks	Document land/water methodology.	Verify and validate methodologies within the context of their use in radiometric corrections. Adopt common methodology and standards for use on multiple sensors.

	Snow/ice masks	Document snow and ice detection methodology.	Verify and validate snow and ice detection methodologies. Adopt common methodology and standards for use on multiple sensors.
	DEM	The required accuracy of the DEM is dependent upon the corrections implemented, swath width, and pixel size.	Share DEMs when appropriate, both among operational agencies and with users. Requirements are highly variable.
	Terrain shadow mask	Terrain shadow masks are needed to estimate radiometric contamination associated with shadows. Known confusion such as with water and cloud shadow needs to be quantified.	Terrain shadows are particularly important for mountainous areas, wide swaths, and for SAR sensors. Adopt common methodology and standards for use on multiple sensors.
	Illumination and viewing geometry	Solar angles are needed for reflectance calculations. View angles are needed for BRDF-related corrections, including terrain illumination corrections.	Establish per-pixel versus scene-center solar angle corrections. View angle corrections.
	Data quality	No data, saturated, contaminated, terrain occlusion pixels need to be identified.	Establish standardised QA mask for different product levels. Adopt common methodology and standards for use on multiple sensors.

Per-pixel metadata are a combination of mask data derived directly from the data and DEM data. The analyses of the items below focuses on the metadata themselves, whereas the use of these data are discussed in the data measures (radiometry) and geometry sections.

Cloud Cover

Cloud cover assessments are derived from the available bands, and the quality of the assessment will vary depending on which bands are available [46, 47]. Cloud models using both thermal and cirrus bands provide the best results. Research has suggested that the inclusion of additional bands in future missions can increase the quality of the cloud cover estimates [48, 49]. The single most influential band is the cirrus band, closely followed by the thermal band. Cirrus cloud estimates over high-elevation land masses may be contaminated by reflectance from the land surface.

The impact of thin clouds on data analysis is application specific. The attenuation of the signal caused by thin clouds must be considered within the overall methodology, total error budget, and specific analysis requirements.

The ability to form dense time series of reflectance data also opens the possibility of multi-temporal cloud masking approaches. For example, the Centre National D'Etudes Spatiales (CNES) MAJA algorithm relies on temporal consistency for identifying clouds [50-53]. These approaches could be extended to multi-sensor datasets.

> *Cloud algorithms need to be verified and validated. Algorithms to minimise variability need to be shared when appropriate. Different band availability will cause differential results. These differences need to be documented. Known confusion, such as with snow and ice, needs to be quantified.*

Cloud Shadow

Cloud shadow models are geometric models relating cloud objects to related dark objects as a function of solar and viewing angles, plus elevation data. Cloud shadows can be confused with terrain shadows, water, and very dark surface features. Cloud shadows contain spectral information that can be used within many applications. Adaptive algorithms can use the masks as information within models.

Cloud shadow models need to be verified and validated. Known confusion, such as with water, needs to be quantified.

Land/Water Masks

Land and water masks generated during the preprocessing of reflectance data are useful as first approximations for atmospheric corrections and higher level product generation. However, these masks are not appropriate for land cover change or water body detection.

Consistently handle differential corrections over land and water.

Snow/Ice Masks

Snow and ice masks created during the preprocessing of sensor data assist in cloud cover assessment. As in the case of land and water masks, snow and ice masks can only serve as first approximations for monitoring, since there is significant confusion with clouds. However, time series of masks can contribute to the detection of clouds and in the monitoring of snow and ice.

Verified and validated uncertainty estimates improve both applications. Known confusion, such as with clouds, needs to be quantified.

Digital Elevation Models (DEMs)

DEMs are, in most cases, external input data used in data production by CEOS agencies. DEMs are required for parallax correction and orthorectification of products [54-56]. DEMs are used in many models, including terrain shadow, geolocation, BRDF, and parallax correction. Some sensors, such as SPOT [57], ASTER [58], ALOS [59], and TanDEM-X [60] can be used to produce DEMs. Many of the publically available DEMs are still based on SRTM data [61]. DEMs are also important inputs for many application models.

SAR data production is particularly sensitive to DEM accuracy and resolution. Sigma naught products are not orthorectified, providing an opportunity to use locally optimised DEM data. Gamma naught products are orthorectified, providing a higher level, analysis-ready data product.

Share DEMs as appropriate both among operational agencies and with users. User requirements for DEMs may differ from data producers and need to be selected to meet sensor-specific requirements. Production and use requirements are highly variable.

Terrain Shadow Masks

Terrain shadows are calculated using solar illumination angles and DEMs. Terrain and cloud shadows both significantly affect radiometry and can cause features in shadow to be confused with other dark features. Terrain shadows contain spectral information that can be used within many applications. Adaptive algorithms can use the masks appropriately.

Terrain shadows are particularly important for mountainous areas, wide swaths, and for SAR sensors. Known confusion, such as with cloud shadow and water, needs to be quantified.

Illumination and Viewing Geometry

Measurement normalisation for solar illumination and viewing angles are critical for BRDF and atmospheric models and become increasingly important for wider swaths. Corrections can include per-pixel solar illumination, viewing angle from sensor, and terrain orientation corrections.

Documented, validated, and verified corrections are critical for time-series analysis using scene overlap regions and multi-sensor datasets.

Data Quality

Data quality metadata need to be well documented and vary significantly between sensors and between products from the same sensor. Quality issues, such as dropped pixels, dropped lines, and saturation, are more common among older sensors, particularly 8-bit sensors. Quality flags also identify no data areas, such as fill, Landsat 7 scan-gaps, and terrain occlusion. If data quality masks are resampled, adjacent pixels can be contaminated and flagged pixels can be lost. Uncertainty estimates associated with data quality should be provided as possible and appropriate.

The distinction between pixels that are contaminated by the attenuated signal caused by reduced optical transparency or shadow versus pixels that need to be flagged as no data is application dependent. Quality assessment information for higher level products may be available to help users make this determination.

> *Define data quality and distinguish between no data and contaminated data. Document how these pixels are handled during resampling.*

Data Measurements

Interoperability concerns include compensation for available bands, normalisation, atmospheric corrections, and spectral band differences for direct measurements from sensors and numeric-derived products such as NDVI. Interoperability of classed data requires knowledge of classification accuracy and confusion among classes within each classification and associations between the product classifications.

Spectral measurements uncertainty will accumulate above and beyond at-sensor noise as the data are resampled and as corrections and application models are applied. Understanding how the uncertainties are distributed for each sensor by processing step helps elucidate the uncertainties that are inherited by any combined multi-sensor dataset.

Perhaps obviously, application models may be sensitive to the existence of specific spectral bands, which may preclude the use of multi-sensor datasets. Spectral band differences need to be accommodated when those differences are a significant proportion of the overall error budget. Spectral band adjustment factors can be used to accommodate band differences, but research is needed to quantify the magnitude of the differences caused by different spectral response curves over different surface types. Indices such as NDVI [41] are sensitive to band differences.

Reflectance calculations at low solar elevation angles can contribute significant uncertainty to measurement estimates. Reflectance is usually not calculated for low sun elevations threshold (below 15-25 degrees), and the threshold will vary by application.

Resampling and geolocation errors also contribute to measurement uncertainty and will be discussed in more detail within the geolocation section below. Table 5 summarises Framework thresholds and targets for items associated with the data measurements component.

Table 5. Data Measurements

Component	Items	Threshold	Target
Data Measurements	Measurements	Document absolutely calibrated measurement. Feasible goals for current missions is 3% at-sensor accuracy and 5.8% at-surface reflectance[49].	Validate and verify surface reflectance data. Feasible goals for future missions is 2% absolute accuracy at-sensor reflectance and 3.6% at-surface reflectance[49].

	Measurement normalisation	Normalise measurements to nadir viewing and temporally constant by spatially varying by latitude solar angle. Use consistent methodology to create multi-sensor data stream.	Investigate more complete, but practical BRDF models, which will require prior knowledge of the Earth surface.
	Aerosol/water vapor/ozone corrections	Document atmospheric model corrections. Use consistent methodology to create multi-sensor data stream.	Validate and verify atmospheric models and compare results. If convergence on single model is not possible, document and accommodate differences.
	Spectral band difference corrections	Initial estimate is a linear fit between equivalent spectral bands using hyperspectral spectra.	SBAFs need to compensate for different spectral response curves, which are surface-type dependent.

The radiometric accuracy has an error budget with contributions from at-sensor within-sensor calibration, at-sensor multi-sensor cross-calibration, spatial misregistration, atmospheric correction, solar angle correction, and view angle correction [41, 62, 63]. Understanding how each of these contributes to the total uncertainty of the estimates is critical to interoperability.

Measurements

Minimum requirement is at-sensor reflectance calibrated and validated to a known absolute source and trended using pseudo-invariant calibration sites. Atmospheric, solar angle, and viewing angle SBAF corrections are usually needed to minimise variability and uncertainty. For higher level products, uncertainties relevant for those products are needed. The next step reduction in uncertainty requires additional spectral bands for atmospheric characterization. The estimated improvement in accuracy is associated with surface reflectance estimates that are on the order of 5% currently to 3% with additional bands [49].

> *Minimum requirement is for optical data that are reflectance calibrated and validated to a known absolute source. Atmospheric, solar angle, and viewing angle SBAF corrections are usually needed to minimise variability and uncertainty to create an interoperable product.*

Measurement Normalisation

The view angle of nadir-pointing moderate resolution data ranges from +/-7.5° for Landsat to +/-10.5° for Sentinel-2 to +/-25° for AWiFS. Even within a sensor record, the view angle for a point on the Earth will vary widely if side lap regions are included in time series. Off-nadir pointable sensors, such as SPOT and ASTER, can have greater view angles even though their field of view is small, resulting in large potential differences in view angle for any given point on the Earth.

For multi-sensor datasets, the view angles for a point on the Earth will be highly variable through time if all observations are used. When solar angles through the year are included, the illumination will also be variable. Measurements can be normalised by adjusting the view angle to nadir and the solar illumination angle to a scene constant as a function of latitude [64-67].

At high latitudes, reflectance products may not be available for part of the year. One assumption for solar angle for the MRI Framework is to only include AM instruments. As the mean crossing time changes through the life of an instrument, uncertainty can be introduced [45].

As a consequence of viewing angle variability, some observations will be toward the sun and others away from the sun and on terrain with different slopes and orientations, thus further complicating models to require knowledge about the Earth’s surface, such as slope, aspect, and surface texture that cannot be directly inferred from the satellite observations. Data can be normalised without introducing full bidirectional reflectance distribution function models; however, surface feature specific variability will remain. [41].

> *Solar and viewing angle corrections compensate for both within- and between-sensor variability. Combining ascending and descending rows adds a further level of complexity. Significant research is needed to establish an optimal and achievable level of correction.*

Aerosol, Water Vapor, and Ozone Corrections

Different atmospheric models can introduce significant between-sensor variability [63, 68]. Atmospheric models must either be shared or validated and verified to a common reference. Available bands and ancillary data will impact corrections; for example, if newer generations of sensors capture atmospheric water, these corrections would be better/more reliable than the older generations where estimates are required. Uncertainty of external atmospheric products will contribute to the overall uncertainty associated with the corrections [62, 64].

> *Atmospheric models continue to improve for moderate resolution satellites. A major challenge is extending the models into the past to include early Landsat and SPOT sensors. Common models should be applied to minimise variability within a single multi-sensor dataset.*

Spectral Band Difference Corrections

Spectral bands can be adjusted using hyperspectral data such as EO-1 Hyperion data [17, 64, 69-76] to determine a regression between similar bands. This mechanism is used to cross-calibrate sensors using spectrally flat pseudo-invariant calibration sites. SBAF becomes a mechanism for homogenization of the similar sensors bands to meet monitoring requirements for specific surface characteristics. Radiometric responses between Sentinel-2 MSI and Landsat 8 OLI can be as high as 17% for different surface types, if SBAFs are not applied [74]. The uncertainty following SBAF correction should approach 5% [77]. If no overlap exists between spectral response curves, no accommodation is possible.

> *SBAFs compensate for different spectral response curves and may be application and surface type specific. Simple solutions are currently available to support current research and applications.*

Geolocation

The analysis of image data through time requires accurate and precise internal geometry and registration to an absolute reference image. Sources of errors lie with the reference database and within each sensor. Within-sensor error components include systematic correction of the data based on ephemeral data acquired with the image and on terrain data. The importance of terrain data is also a function of path width, with wider paths being more sensitive to terrain. The systematic and terrain corrected information is convolved with a reference grid accuracy to provide a measure of absolute geometry error and relative error among sensors. Reference grids are created and validated using independent, high-resolution data. A shared reference grid and a high-quality DEM are major steps to achieving geometric interoperability between sensors.

The study of geolocation error will be coordinated through WGCV and at individual CEOS agencies. Several studies of note are addressing geolocation uncertainty from the perspective of uncertainty in relation to ground reference, both within sensor and between sensors [34].

Interoperability issues include compensation for different pixel sizes between products, spatial RMSE by image, reference grids with different absolute accuracies, DEMs with different accuracies, and different projection spaces. Table 6 summarises Framework thresholds and target for items associated with the geolocation component.

Table 6. Geolocation

Component	Items	Threshold	Target
Geolocation	Geometric correction	Image data are precision-corrected to a reference dataset.	Minimise misregistration through orthorectification and precision registration to a common reference dataset. Document methods and uncertainties/error throughout the processing chain.
	Resampling	The number and type of spatial resampling will impact the radiometric signal.	Minimise the number of resampling operations. Quantify impact of upsampling or downsampling on time-series analysis for different applications. Document resampling type/method applied.

Geometric Correction

Misregistration between images effectively increases variability in the radiometry measurements. Misregistration is minimised through orthorectification and precision registration to a controlled reference grid. Table 7 summarises variables that contribute to misregistration. The rule of thumb for interoperability is sub-pixel accuracy with acknowledgement that any misregistration adds noise to time-series applications and each application will have different tolerances for the increased noise. The geometric accuracy required for time-series analysis depends on the application. It is understood that radiometric noise is introduced into the time series as a function of the pixel size, point spread function, and misregistration. User guidelines are needed to support application specific decisions. Acceptable accuracy will vary from on the order of one pixel to less than 0.1 pixel depending on the sensitivity of the monitored change. These rules of thumb will vary depending on the size and contrast of features as well as the point spread function of the sensor. Practical guidelines depend on how well different sensors can be registered and combined with available sensors that meet application location and date requirements (Table 7). Users will accommodate the best available data and adapt analytic methodologies and goals based on what is achievable.

Example measured difference between Landsat 8 and Sentinel-2 before convergence on the Sentinel-2 GRI is 26 meters, which must be corrected. When Landsat changes to using the GRI, then the result should approach 10m. Current approaches require an extra resampling until the change is made.

> *Verify image-to-image registration and remove unacceptable misregistration. Minimise misregistration through orthorectification and precision registration to a controlled and preferably shared reference grid.*

Table 7. Geolocation Characteristics by Sensor

Platform	Instrument	Reference Grid	DEM	Pixel Size	RMSE	References
Landsat	MSS	GLS 2000 (17m)	GLS DEM	79m	134m CE95	[5, 78]
	TM	GLS 2000 (17m)	GLS DEM	30m	10.9m CE95	[78]
	ETM+	GLS 2000 (17m)	GLS DEM	30m	10.7m CE95	[78]
	OLI/TIRS	GLS 2000 (17m)	GLS DEM	30m	15m CE95	[34, 78-80]
Sentinel-2	MSI	GRI (9.5m)	PlanetDEM	20/30m	12.5m CE95	[34, 36, 37]

TERRA	ASTER	GLS 2000 (17m)				
SPOT	HRV, HRG					
CBERS	MUX, WFI, IRS					
ResourceSat	AWIFS, LISS-III			56m		[81-85]
EO-1	Hyperion, ALI			30m		[16, 17]
Sentinel-1	C-band					
Radarsat 2	C-band					
ALOS-1/2	PALSAR L-band					
TerraSAR-X **TanDEM-X**	X-band					

Resampling

To homogenize sensor products to meet RMSE requirements or pixel size using image-to-image registration requires an additional resampling of the data. Upsampling data to smaller pixel size will retain the resolution of the higher resolution sensor at the cost of introducing increased radiometric variability, since upsampled data imputes information for the lower resolution sensor as a function of the resampling mechanism (e.g., nearest neighbor, bilinear, or cubic convolution) used. Downsampling data results in the loss of spatial resolution, but creates a single dataset of consistent radiometry. Resampling of classed products can introduce significant new uncertainty in the products.

The identification of a shared reference grid is an important first step for achieving geolocation interoperability. To minimise impact on radiometry, resampling must be minimised, particularly if the pixel size is changed. The adoption of a shared mapping grid for high-level products, such as the Discrete Global Grid Systems (DGGS) is an opportunity to reduce resampling. Resampling to a shared projection space should carried out as early in the production flow as possible, preferably while the data are still path aligned.

> *Minimise the number of resampling operations and consider impacts such as contaminated pixels, clouds, and the resampling method.*

Case Studies

To date we have initiated one case study based upon the NASA Harmonized Landsat/Sentinel-2 (HLS) Project (see Appendix B). This product was characterised and put into the MRI Framework. Although this example is one of backing an existing merged product into the Framework rather than building requirements from each of the contributing sensors, it is informational in that it is a clear example of a successful exploration of the interoperability of these two key sensors.

Other case studies are planned for 2018 and beyond, an effort that will be led by LSI-VC as an ongoing task. The first of those case studies will be a user-based project focused on vegetation dynamics monitoring with harmonized Landsat 8 and Sentinel-2 data. This case study utilises the HLS 30m products at several global locations to determine whether it is possible to detect and monitor vegetation productive dynamics and phenology (i.e., reliably pick the seasonal cycles) based on these and other ancillary data.

Interoperability research and applications are increasingly common throughout the user community. These serve as a rich source of case studies covering a wide range of lessons learned and good practices for many sensors and thematic areas. A survey was designed by the MRI team and was initiated with the hope that it will eventually be implemented to identify case studies, lessons learned, and good practices from throughout the user community. The survey will be completed in 2018.

Conclusions

The USGS 2017 Chair Moderate Resolution Sensor Interoperability (MRI) initiative bridges the gap between CEOS data products and the user community for the implementation of consistent and complementary multi-sensor data streams. With the current suite of free and open access data products, the development of long-term (1972 to the present) and dense (2-4 day revisit) time series are possible, if the products are, or can be adapted to be, interoperable. These time series can be used to implement a new generation of analysis and monitoring methodologies. The Future Data Architecture (FDA) Initiative can take advantage of the lessons learned and good practices identified to implement these new methodologies for use by the thematic communities.

Through the continued evolution of the MRI Framework, higher level data products and SAR data can be more completely integrated. New case studies can continue to build on the compilation of lessons learned and best practices needed by the user community. Going forward, relationships are needed between the MRI initiative and R&D and capacity-building teams to understand and communicate the benefits and challenges of multi-sensor interoperability for user communities.

Acknowledgements

Members of the 2017 MRI initiative team: Gene Fosnight (USGS), Richard Moreno (CNES), Cindy Ong (CSIRO), Jeff Masek (NASA), Zoltan Szantoi (EC), Andy Mitchell (NASA), Kurt Thome (NASA), Kevin Gallo (NOAA), Kerry Sawyer (NOAA), Paul Briand (CSA), Yves Crevier (CSA), Brian Killough (NASA), Nigel Fox (UKSA), Amanda Regan (EC), Debajyoti Dhar (ISRO), Takeo Tadono (JAXA), Koji Akiyama (JAXA), Adam Lewis (GA), Matt Steventon (Symbios), Eric Wood (USGS)

Appendix A: CEOS 10-100 Meter Resolution Sensor Characteristics

The table below provides a high-level summary of many CEOS agency moderate resolution (10-100m) optical and SAR sensors [3, 4].

	Platform	Instrument	Radiometry	Pixel Size	FOV	Life	Reference
Optical	Landsat	MSS	Visible, NIR	79m	15˚	1972-2008	[5-7]
		TM, ETM+	Visible, NIR, SWIR, Thermal	30m	15˚	1982	[8, 9]
		OLI/TIRS	Visible, NIR, SWIR, Cirrus, Thermal	30m	15˚	2013	[10]
	Sentinel-2	MSI	Visible, NIR, SWIR, Cirrus	10/20/60m	21˚	2015	
	Terra	ASTER	Visible, NIR, SWIR, Thermal	15/30/90m		1999	[11]
	SPOT	HRV	Visible, NIR	20m			[12, 13]
		HRVIR, HRG	Visible, NIR, SWIR	10/20m			
	CBERS	MUX, WFI, IRS	Visible, SWIR, NIR, Thermal	20/40/64/80m			[14]
	ResourceSAT	AWIFS, LISS-III	Visible, NIR, SWIR	56/23.5m	50˚/15˚		[15]
	EO-1	Hyperion	Visible, NIR, SWIR				[16, 17]
SAR	Sentinel-1	C-band				2016	
	Radarsat 2	C-band					
	JERS-1	L-band	HH polarization	25m		1993-1998	[18]
	ALOS-1/2	L-band	HH+HV	25m		2007	[18]
	TerraSAR-X TanDEM-X	X-band					

Appendix B: Harmonized Landsat/Sentinel-2 (HLS) Case Study

This case study can serve as an example of the use of the MRI Framework for attaining multi-sensor interoperability. The objective of the NASA Harmonized Landsat/Sentinel-2 (HLS) Project is to generate a radiometrically and geometrically (seamless and interchangeable) surface reflectance dataset from Landsat 8 and Sentinel-2 [64, 77]. Each observation is manipulated to look like it came from a single sensing system with the consistent statistical properties, and the origin of each observation is transparent to end users. HLS performs a series of consistent radiometric and geometric corrections to minimise sensor differences, including a common atmospheric correction, solar/view angle corrections, spectral bandpass adjustments, and gridding to a common UTM projection and tiling system.

Component	**Items**	**Descriptions**
General Metadata	Coordinate reference system	Both products are projected to a UTM/WGS84 map projection. HLS merged products are produced as 30-meter products using the Sentinel-2 tile system. Individual products are available at native resolution prior to resampling Sentinel-2 to 30 meters.
	Reference grid accuracy	Landsat data are georegistered to the GLS reference database, which contains errors of up to 36m. In order to provide consistent georeferencing, both Landsat 8 and early (pre-v2.04 processing system) Sentinel-2 data are registered and resampled to the best available Sentinel-2 MSI image. L30 products have absolute geodetic error of <19m (CE90), while S30 products have absolute geodetic error <10.5m (2σ).
	Geometric accuracy and temporal consistency	L30 products have absolute geodetic error of <19m (CE90), while S30 products have absolute geodetic error <10.5m (2σ).
	Spectral bands	The common bands are coastal aerosol, blue, green, red, NIR, SWIR1, and SWIR2. Other bands are available for analysis.
	Spectral response curves	The Sentinel-2 MSI band passes are quite similar to those of Landsat 8 OLI, for those bands in common between both instruments. The near-infrared (MSI Band 8a) and shortwave bands in particular are nearly identical. The MSI green band is slightly broader in comparison to OLI, while the red band is shifted ~15nm to the shorter end of the spectrum. HLS uses a linear regression model (based on training sample from Hyperion hyperspectral data) to derive Spectral Band Adjustment Factors (SBAF) that convert Sentinel-2 reflectances into 'equivalent' Landsat OLI reflectances for the common bands.
	Radiometric accuracy	Sentinel-2 MSI has a radiometric stability (i.e., uniform target over multiple overpasses) requirement of better than 1% (2σ) while Landsat 8 OLI has a requirement of better than 0.5% (2σ). MSI and OLI agree to within 1.5% with the exception of the coastal aerosol and blue bands.
	Revisit time & lifetime	Landsat 8 OLI revisit time is every 16-days and Sentinel-2 A & B revisit every 5 days over Greenland, Europe, and Africa and every 10 days over the rest of the world, as of July 2017. Given swath overlap and average cloud cover, a cloud-free HLS observation (either L30 or S30) can be expected every 15-20 days over the humid tropics, and every 5-10 days over mid-latitude agricultural regions.
	Field of view	The Landsat 8 FOV is 15° and the Sentinel-2 FOV is 21°. View angle differences for some ground targets can differ by up to 18.5°.
	Mean local time	The average mean local time for Landsat 8 is 10:11 and for Sentinel-2 is 10:30. There is minimal impact on radiometry.
Per-pixel Metadata	Cloud cover	The FMask algorithm is used for both Landsat 8 and Sentinel-2 to detect clouds. The lack of a thermal band on Sentinel-2 increases errors of omission and commission.
	Cloud shadow	The FMask algorithm is used for both Landsat 8 and Sentinel-2 to detect cloud shadows. Adjacent cloud pixels are only estimated for Landsat 8.
	Land/water masks	The FMask algorithm is used for both Landsat 8 and Sentinel-2 to detect water.

Component	Items	Descriptions
	Snow/ice masks	The FMask algorithm is used for both Landsat 8 and Sentinel-2 to detect snow and ice.
	DEM	Landsat uses the GLS DEM. Sentinel-2 uses the PlanetDEM.
	Terrain shadow mask	Not used in HLS products.
	Illumination and viewing geometry	Solar illumination angles are needed for reflectance calculations. Solar and view angles are needed for BRDF related corrections. View geometry is provided on a per-pixel basis for both Landsat 8 and Sentinel-2 data.
	Data quality	HLS includes Quality Assessment on a per-tile and per-granule basis, by comparison with contemporary MODIS CMG (Climate Modeling Grid) BRDF-adjusted reflectances. QA summaries are available on the HLS web site.
Data Measurements	Measurements	HLS products record surface reflectance or apparent (TOA, blackbody) temperature.
	Measurement normalisation	Reflectance values are normalised to a constant (nadir) view angle and fixed, latitude-dependent solar elevation using the coefficients provided in Roy, et al. [66].
	Aerosol/water vapor/ozone corrections	Aerosol quality flags are set for Landsat 8. Cirrus per-pixel flags are set for both.
	Spectral band difference corrections	Sentinel-2 reflectance values are adjusted to match those derived from equivalent Landsat 8 bandpasses, using a linear-regression model trained on Hyperion hyperspectral data.
Geolocation	Geometric corrections	The HLS product uses image-to-image registration to a single reference Sentinel-2 image for coregistration of each tile to compensate for different reference databases [77].
	Resampling	Sentinel-2 data are resampled to 30 meters to preserve the radiometric time series at the cost of lower spatial resolution.

Foundational efforts are already underway to ensure interoperability [64, 86-89], and include pre-flight and on-orbit cross-calibration of Sentinel-2 carried out by NASA, USGS, and ESA. As Sentinel-2's product generation pipelines are fully implemented, a need for higher level coordination exists.

References

1. CEOS LSI-VC, *CEOS Analysis Ready Data for Land (CARD4L) Description Document - Version 22*. 2016, CEOS LSI-VC. p. 2.
2. CEOS FDA Team, *CEOS Future Data Access & Analysis Architectures Study Interim Report Version 1.0 - October 2016*. 2016.
3. Gómez, C., J.C. White, and M.A. Wulder, *Optical remotely sensed time series data for land cover classification: A review.* ISPRS Journal of Photogrammetry and Remote Sensing, 2016. **116**: p. 55-72.
4. CEOS. *CEOS Missions, Instruments and Measurements Database Online - updated for 2017*. 2017 [cited 2017 13 February 2017]; available from http://database.eohandbook.com/.
5. USGS EROS, *MSS Image Assessment System: Geometric Algorithm Description Document*. 2011. p. 51.
6. USGS EROS, *Landsat 1/5 Multispectral Scanner (MSS) Image Assessment System (IAS) Radiometric Algorithm Description Document (ADD)*. 2011, Department of the Interior/USGS. p. 84.
7. USGS EROS, *Landsat Multispectral Scanner (MSS) Level 1 (L1) Data Format Control Book (DFCB) - Version 9.0*. 2015, USGS EROS: Sioux Falls, SD, USA. p. 36.
8. USGS EROS, *Landsat 7 ETM+ Level 1 Product Data Format Control Book (DFCB) - Version 18.0*. 2015, USGS EROS: USGS EROS, Landsat Project, Sioux Falls, SD, USA. p. 36.
9. USGS EROS, *Landsat Thematic Mapper (TM) Level 1 (L1) Data Format Control Book (DFCB) - Version 8.0*. 2015, USGS EORS: Sioux Falls, SD, USA.
10. USGS EROS, *Landsat 8 Level 1 Data Format Control Book (DFCB) - Version 9.0*. 2015, USGS EROS: Sioux Falls, SD, USA.
11. Duda, K. and J. D'aucsavage, *Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Level 1 Precision Terrain Corrected Registered At-Sensor Radiance Product (AST_L1T)* 2015.
12. Riazanoff, S., *SPOT 123-4-5 Geometry Handbook*. 2004.
13. SPOT IMAGE, *SPOT satellite technical data*. 2010.
14. Pinto, C., et al., *First in-Flight Radiometric Calibration of MUX and WFI on-Board CBERS-4.* Remote Sensing, 2016. **8**(5).
15. NRSC, *Resourcesat-2*. n.d., National Remote Sensing Centre, Indian Space Research Organisation, Dept. of Space, Govt. of India. p. 8.
16. Bridgman, T., S. Ungar, and L. Ong. *Comparing EO-1-Hyperions spectral resolution to Landsat*. [Internet Resource; Visual Material] 2001; available from http://svs.gsfc.nasa.gov/vis/a000000/a002000/a002097/index.html.
17. Chander, G., et al. *Use of EO-1 hyperion data to calculate spectral band adjustment factors (SBAF) between the L7 ETM+ and terra MODIS sensors*. in *International Geoscience and Remote Sensing Symposium (IGARSS)*. 2010.
18. JAXA, *Global 25m Resolution PALSAR-2/PALSAR Mosaic and Forest/Non-Forest Map (FNF) Dataset Description* 2017, Japan Aerospace Exploration Agency (JAXA) Earth Observation Research Center (EORC). p. 9.

19. Hostert, P., A. Röder, and J. Hill, *Coupling spectral unmixing and trend analysis for monitoring of long-term vegetation dynamics in Mediterranean rangelands.* Remote Sensing of Environment, 2003. **87**(2-3): p. 183-197.

20. Kennedy, R.E., Z. Yang, and W.B. Cohen, *Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms.* Remote Sensing of Environment, 2010. **114**(12): p. 2897-2910.

21. Senf, C., et al., *A Bayesian hierarchical model for estimating spatial and temporal variation in vegetation phenology from Landsat time series.* Remote Sensing of Environment, 2017. **194**: p. 155-160.

22. Verbesselt, J., et al., *Detecting trend and seasonal changes in satellite image time series.* Remote Sensing of Environment, 2010. **114**(1): p. 106-115.

23. Verbesselt, J., et al., *Phenological change detection while accounting for abrupt and gradual trends in satellite image time series.* Remote Sensing of Environment, 2010. **114**(12): p. 2970-2980.

24. Verbesselt, J., A. Zeileis, and M. Herold, *Near real-time disturbance detection using satellite image time series.* Remote Sensing of Environment, 2012. **123**: p. 98-108.

25. Zhu, Z. and C.E. Woodcock, *Continuous change detection and classification of land cover using all available Landsat data.* Remote Sensing of Environment, 2014. **144**: p. 152-171.

26. Cohen, W.B., Z. Yang, and R. Kennedy, *Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. TimeSync — Tools for calibration and validation.* Remote Sensing of Environment, 2010. **114**(12): p. 2911-2924.

27. Hamunyela, E., J. Verbesselt, and M. Herold, *Using spatial context to improve early detection of deforestation from Landsat time series.* Remote Sensing of Environment, 2016. **172**: p. 126-138.

28. CEOS LSI-VC, *CEOS Analysis Ready Data for Land (CARD4L) Specification Framework v1.3.0* A.s.f. v1.3.0.xlsx, Editor. 2017.

29. CEOS LSI-VC, *CARD4L Product Family Specification - Surface Termperature (CARD4L-ST)*. 2017. p. 10.

30. CEOS LSI-VC, *CARD4L Product Family Specification - Optical Surface Reflectance (CARD4L-OSR)*. 2017.

31. CEOS LSI-VC, *CARD4L Product Family Specification - Normalised Radar Backscatter (CARD4L-NRB)*. 2017.

32. Purss, M.B.J., et al., *Discrete Global Grid Systems for Handling Big Data from Space*. 2016, Open Geospatial Consortium Discrete Global Grid Systems Standards Working Group.

33. Rengarajan, R., et al., *Validation of geometric accuracy of global land survey (GLS) 2000 data.* Photogrammetric Engineering and Remote Sensing, 2015. **81**(2): p. 131-141.

34. Storey, J., et al., *A note on the temporary misregistration of Landsat-8 Operational Land Imager (OLI) and Sentinel-2 Multi Spectral Instrument (MSI) imagery.* Remote Sensing of Environment, 2016. **186**: p. 121-122.

35. Déchoz, C., et al., *Sentinel-2 Global Reference Image*, in *Image and Signal Processing for Remote Sensing XXI*, L. Bruzone, Editor. 2015.

36. Bruzzone, L., et al., *Sentinel-2: presentation of the CAL/VAL commissioning phase.* 2015. **9643**: p. 964309.

37. Gaudel, A., et al., *Sentinel-2 Global Reference Image Validation and Application to Multitemporal Performances and High Latitude Digital Surface Model.* ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2017. **XLII-1/W1**: p. 447-454.

38. Storey, J. and M. Choate. *Landsat 7 on-orbit geometric calibration and performance*. 2000.

39. Storey, J., M. Choate, and K. Lee, *Landsat 8 operational land imager on-orbit geometric calibration and performance.* Remote Sensing, 2014. **6**(11): p. 11127-11152.

40. Yan, L., et al., *An Automated Approach for Sub-Pixel Registration of Landsat-8 Operational Land Imager (OLI) and Sentinel-2 Multi Spectral Instrument (MSI) Imagery.* Remote Sensing, 2016. **8**(6): p. 520.

41. Roy, D.P., et al., *Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity.* Remote Sensing of Environment, 2016. **185**: p. 57-70.

42. Tian, F., et al., *Evaluating temporal consistency of long-term global NDVI datasets for trend analysis.* Remote Sensing of Environment, 2015. **163**: p. 326-340.

43. Barsi, J., et al., *Sentinel-2A MSI and Landsat-8 OLI Radiometric Cross Comparison.* submitted to European Journal of Remote Sensing, 2017: p. submitted.

44. Whitcraft, A.K., et al., *Cloud cover throughout the agricultural growing season: Impacts on passive optical earth observations.* Remote Sensing of Environment, 2015. **156**: p. 438-447.

45. Zhang, H.K. and D.P. Roy, *Landsat 5 Thematic Mapper reflectance and NDVI 27-year time series inconsistencies due to satellite orbit change.* Remote Sensing of Environment, 2016. **186**: p. 217-233.

46. Hollstein, A., et al., *Ready-to-Use Methods for the Detection of Clouds, Cirrus, Snow, Shadow, Water and Clear Sky Pixels in Sentinel-2 MSI Images.* Remote Sensing, 2016. **8**(8): p. 666.

47. Zhu, Z., S. Wang, and C.E. Woodcock, *Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images.* Remote Sensing of Environment, 2015. **159**: p. 269-277.

48. Landsat Science Team. *Landsat 10 Requirements and Capabilities - team member perspectives on future needs*. in *Landsat Science Team Meeting*. 2016. Brookings, SD.

49. Helder, D. *EROS Cal/Val Center of Excellence*. in *Landsat Science Team Meeting*. 2017. Sioux Falls, SD.

50. Mira, M., et al., *Uncertainty assessment of surface net radiation derived from Landsat images.* Remote Sensing of Environment, 2016. **175**: p. 251-270.

51. Hagolle, O., et al., *A Multi-Temporal and Multi-Spectral Method to Estimate Aerosol Optical Thickness over Land, for the Atmospheric Correction of FormoSat-2, LandSat, VENS and Sentinel-2 Images.* Remote Sensing, 2015. **7**(3): p. 2668-2691.

52. Hagolle, O., et al., *A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENμS, LANDSAT and SENTINEL-2 images.* Remote Sensing of Environment, 2010. **114**(8): p. 1747-1755.

53. Hagolle, O., et al., *Correction of aerosol effects on multi-temporal images acquired with constant viewing angles: Application to Formosat-2 images.* Remote Sensing of Environment, 2008. **112**(4): p. 1689-1701.

54. Ressl, C. and N. Pfeifer, *Evaluation of the Terrain Modl Influence on the Orthorectification of*

Sentinel-2 Satellite Images Over Different Land Forms in Austria. n.d.

55. Storey, J., *Landsat Terrain Error Sensitivity Study – Version 1.1*. 2003: U.S. Geological Survey: Sioux Falls, SD, USA, 2003.

56. Storey, J., M. Choate, and K. Lee. *Geometric performance comparison between the OLI and the ETM+*. in *Pecora 17 - The Future of Land Imaging…Going Operational*. 2008. Denver, CO.

57. SPOT IMAGE, *SPOT DEM Precision Product description Version 1.0*. 2006. p. 18.

58. Tachikawa, T., et al., *ASTER Global Digital Elevation Model Version 2 - Summary of Validation Results*. 2011.

59. JAXA, *ALOS Global Digital Surface Model (DSM) "ALOS World 3D-30m" (AW3D30) Data set - Product Format Description Version 1.1*. 2017, Earth Observation Research Center (EORC), Japan Aerospace Exploration Agency (JAXA). p. 8.

60. DLR, *TanDEM-X Ground Segment DEM Products Specification Document*. 2013, German Aerospace Center (DLR) Earth Observation Center. p. 43.

61. Crippen, R.E., et al., *NASADEM GLOBAL ELEVATION MODEL: METHODS AND PROGRESS* in *Photogrammetry, Remote Sensing and Spatial Information Sciences, XXIII ISPRS Congress*. 2016, ISPRS: Prague, Czech Republic.

62. Claverie, M., et al., *Evaluation of the Landsat-5 TM and Landsat-7 ETM+ surface reflectance products.* Remote Sensing of Environment, 2015. **169**: p. 390-403.

63. Vermote, E., et al., *Preliminary analysis of the performance of the Landsat 8/OLI land surface reflectance product.* Remote Sensing of Environment, 2016. **185**: p. 46-56.

64. Claverie, M., J.G. Masek, and J. Ju, *Harmonized Landsat-8 Sentinel-2 (HLS) Product User's Guide*. 2016. p. 16.

65. Li, F., et al., *A physics-based atmospheric and BRDF correction for Landsat data over mountainous terrain.* Remote Sensing of Environment, 2012. **124**: p. 756-770.

66. Roy, D.P., et al., *A general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance.* Remote Sensing of Environment, 2016. **176**: p. 255-271.

67. Zhang, H.K., D.P. Roy, and V. Kovalskyy, *Optimal Solar Geometry Definition for Global Long-Term Landsat Time-Series Bidirectional Reflectance Normalization.* IEEE Transactions on Geoscience and Remote Sensing, 2016. **54**(3): p. 1410-1418.

68. Vermote, E., et al. *CEOS-WGCV Atmospheric Correction Inter-comparison eXercise*. in *1st Workshop of CEOS-WGCV Atmospheric Correction Inter-comparison Exercise*. 2016. ESA/ESRIN, Frascati, Italy.

69. Chander, G., et al., *Applications of Spectral Band Adjustment Factors (SBAF) for Cross-Calibration.* Ieee Transactions on Geoscience and Remote Sensing, 2013. **51**(3): p. 1267-1281.

70. Pinto, C.T., et al., *Evaluation of the uncertainty in the spectral band adjustment factor (SBAF) for cross-calibration using Monte Carlo simulation.* Remote Sensing Letters, 2016. **7**(9): p. 837-846.

71. Chander, G., et al., *Radiometric cross-calibration of EO-1 ali with L7 ETM+ and terra MODIS sensors using near-simultaneous desert observations.* IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2013. **6**(2): p. 386-399.

72. Chander, G., et al., *Assessment of Spectral, Misregistration, and Spatial Uncertainties Inherent in the Cross-Calibration Study.* IEEE Transactions on Geoscience and Remote Sensing, 2013. **51**(3):

p. 1282-1296.

73. Kaewmanee, M., et al. *Inter-comparison of Theos and Landsat-5 TM over the Libya 4 pseudo-invariant calibration site*. 2012.

74. Micijevic, E. and M.O. Haque, *Effects of Spectral Band Differences between Landsat 8 Operational Land Imager (OLI) and Sentinel 2A Multispectral Instrument (MSI)*, in *2015 AGU Fall Meeting*. 2016.

75. Henry, P., et al., *Assessment of Spectral Band Impact on Intercalibration Over Desert Sites Using Simulation Based on EO-1 Hyperion Data.* IEEE Transactions on Geoscience and Remote Sensing, 2013. **51**(3): p. 1297-1308.

76. Chander, G., et al., *Overview of intercalibration of satellite instruments.* IEEE Transactions on Geoscience and Remote Sensing, 2013. **51**(3): p. 1056-1080.

77. Claverie, M., et al., *The Harmonized Landsat and Sentinel-2 data set.* Submitted to Remote Sensing of Environment, 2017.

78. Storey, J., et al., *Sentinel-2 Multi Spectral Instrument (MSI) and Landsat-8 Operational Land Imager (OLI) Registration Accuracy*. 2016.

79. Storey, J., M. Choate, and D. Moe, *Landsat 8 thermal infrared sensor geometric characterization and calibration.* Remote Sensing, 2014. **6**(11): p. 11153-11181.

80. Storey, J.C. and M. Choate, *Geometric and spatial performance of Landsat 8*. 2014, Falls Church, VA: Photogrammetric Engineering and Remote Sensing. 16.

81. Chander, G. and P.L. Scaramuzza. *Cross-calibration of the Landsat-7 ETM+ and Landsat-5 TM with the ResourceSat-1 (IRS-P6) AWiFS and LISS-III sensors*. 2006.

82. Chander, G., et al. *Cross-comparison of the IRS-P6 AWiFS sensor with the L5 TM, L7 ETM+, & Terra MODIS sensors*. 2009.

83. Goward, S.N., et al., *Complementarity of ResourceSat-1 AWiFS and Landsat TM/ETM+ sensors.* Remote Sensing of Environment, 2012. **123**: p. 41-56.

84. Misra, I., et al., *An efficient algorithm for automatic fusion of RISAT-1 SAR data and Resourcesat-2 Optical images*, in *IEEE Proceedings of 4th International Conference on Intelligent Human Computer Interaction*. 2012: Kharagpur, India.

85. Chen, X., et al., *Cross-sensor comparisons between Landsat 5 TM and IRS-P6 AWiFS and disturbance detection using integrated Landsat and AWiFS time-series images.* International Journal of Remote Sensing, 2013. **34**(7): p. 2432-2453.

86. Doxani, G., et al., *A Spectral Unmixing Model for the Integration of Multi-Sensor Imagery: A Tool to Generate Consistent Time Series Data.* Remote Sensing, 2015. **7**(10).

87. Masek, J., et al., *Harmonizing Landsat and Sentinel-2*. 2016.

88. Masek, J., et al., *Harmonized Landsat and Sentinel-2 (HLS) Products Update*, in *Landsat Science Team Meeting*. 2017: Boston.

89. NASA. *Harmonized Landsat Sentinel-2* 2017 [cited 2017 2017-01-20]; available from http://hls.gsfc.nasa.gov/.