



Multispectral high resolution sensor fusion for smoothing and gap-filling in the cloud



Álvaro Moreno-Martínez^{a,b,*}, Emma Izquierdo-Verdiguier^c, Marco P. Maneta^{d,e}, Gustau Camps-Valls^a, Nathaniel Robinson^f, Jordi Muñoz-Marí^a, Fernando Sedano^g, Nicholas Clinton^h, Steven W. Running^b

^a Image Processing Laboratory (IPL), Universitat de València, València, Spain

^b Numerical Terradynamic Simulation Group (NTSG), WA Franke College of Forestry and Conservation, University of Montana, Missoula, USA

^c Institute of Geomatics, University of Natural Resources and Life Sciences, Wien, Austria

^d Department of Geosciences, University of Montana, USA

^e Department of Ecosystem and Conservation Sciences, WA Franke College of Forestry and Conservation, University of Montana, USA

^f Panthera, New York, NY, USA

^g Department of Geographical Sciences, University of Maryland, College Park, USA

^h Google, Inc., Mountain View, CA, USA

ARTICLE INFO

Keywords:

Landsat
MODIS
Gap filling
Smoothing
Kalman filter
Data fusion

ABSTRACT

Remote sensing optical sensors onboard operational satellites cannot have high spectral, spatial and temporal resolutions simultaneously. In addition, clouds and aerosols can adversely affect the signal contaminating the land surface observations. We present a HIghly Scalable Temporal Adaptive Reflectance Fusion Model (HIST-ARFM) algorithm to combine multispectral images of different sensors to reduce noise and produce monthly gap free high resolution (30 m) observations over land. Our approach uses images from the Landsat (30 m spatial resolution and 16 day revisit cycle) and the MODIS missions, both from Terra and Aqua platforms (500 m spatial resolution and daily revisit cycle). We implement a bias-aware Kalman filter method in the Google Earth Engine (GEE) platform to obtain fused images at the Landsat spatial-resolution. The added bias correction in the Kalman filter estimates accounts for the fact that both model and observation errors are temporally auto-correlated and may have a non-zero mean. This approach also enables reliable estimation of the uncertainty associated with the final reflectance estimates, allowing for error propagation analyses in higher level remote sensing products. Quantitative and qualitative evaluations of the generated products through comparison with other state-of-the-art methods confirm the validity of the approach, and open the door to operational applications at enhanced spatio-temporal resolutions at broad continental scales.

1. Introduction

Monitoring of the Earth system with remote sensors has enabled, for the first time in human history, critical information to better study and understand the atmosphere, oceans, land, biosphere, cryosphere as a whole revealing the importance of this information for monitoring patterns and processes which define the Earth's land areas (Goward et al., 2001). Consequently, the broad spectrum of applications of land observation with remote sensing data is continuously growing and spans many fields from disaster management, flood mapping, vegetation monitoring, land surface classification, and the estimation of mass and energy fluxes. Because of the variety of applications, satellite sensors are designed to provide optimal data for their intended

applications. Unfortunately, this specificity hampers the development of general applications and pushes us to integrate the information from different sensors (Thenkabail et al., 2018).

Depending on the application and the spectral region that a remote sensor measures, the availability of spatially and temporally continuous data could be challenging to obtain. Optical remote sensing data, from sensors measuring reflectances in the visible, near-infrared (NIR), and short wave infrared (SWIR) wavelengths, are the most extensively used by the scientific community due to a wide range of applications (Maly et al., 2019). Although there is an extensive array of optical remote sensing data, from a variety of satellites representing a substantial time series, these data are incapable of retrieving reliable land surface information when clouds, aerosols, shadows, and strong angular effects

* Corresponding author.

E-mail address: alvaro.moreno@uv.es (Á. Moreno-Martínez).

are present in the scenes. To alleviate some of these problems, different atmospheric and directions effects correction methods have been proposed and developed in the last decade (Schaaf et al., 2002; Vermote et al., 2002; Lyapustin et al., 2008; Schmidt et al., 2013).

The mitigation of noise and gap filling of satellite data are preliminary tasks for any remote sensing application aimed at effectively analyzing the earth's surface through time. The scientific literature is rich in methods tackling these issues, and solutions vary significantly in terms of complexity and sophistication (Kandasamy et al., 2013; Shen et al., 2015). These solutions are usually classified in different groups: temporal, spatial, spatio-temporal, and spatio-temporal fusion approaches. Temporal methods include thresholding, rank-based (e.g., mean, median, extreme filters), and polynomial (e.g., moving average, Savitzky-Golay, locally weighted scatterplot smoothing) approaches (Jonsson and Eklundh, 2002; Chen et al., 2004; Moreno et al., 2014; Robinson et al., 2017). Most often, local temporal methods do not account for specific processes causing undesired variability in the data in order to reconstruct the time series. They infer the correct or missing information from close observations in the time domain. On the other hand, global temporal approaches fit predefined functions (polynomial, Double Logistic, asymmetric Gaussian) or decomposition techniques based on Fourier analysis (Jonsson and Eklundh, 2002; Bradley et al., 2007) to the data. Because global temporal methods assume a predefined phenological shape, they generally lead to a lack of flexibility, especially when fitted to an irregular or asymmetric time series (Chen et al., 2004).

Spatial gap filling approaches, such as kriging, generally utilize information from non-gap neighboring pixels to infer the missing data. Their main assumption is that the missing information could be retrieved from the spatial autocorrelation structure in the variogram (Matheron, 1963). Improved versions of these spatial methods include a second source of ancillary gap-free observations for the same sensor in a second date as a covariate (Zhang et al., 2007).

Spatio-temporal gap-filling algorithms are generally implemented as multi-step approaches whereby gaps are filled through a series of alternating spatial and temporal steps (Kang et al., 2005). Other more recent approaches utilize Generalized Additive Models (GAM) or extended Singular Spectrum Analysis (SSA) to perform spatio-temporal gap-filling to a higher degree of sophistication (Poggio et al., 2012; Zschieschler et al., 2014).

The increasing number of Earth Observation Satellites (EOS) provides a unique opportunity to merge the information from different sensors. Thus, exploiting synergistically their different angular, spectral, spatial, and temporal sensing characteristics is proving to be an efficient way to overcome individual limitations of each sensor through providing harmonized multi-sensor observations. As an example, the Landsat platforms with their Thematic Mappers (Landsat 4–5), the Enhanced Thematic Mapper plus (Landsat 7), and the Operational Land Imager (Landsat 8) sensors have been widely used in many Earth observation and monitoring applications over the last few decades. Although they have a high spatial resolution (15 to 60 m spatial resolution depending on the sensor), they have a low temporal resolution (16 to 18 days revisit cycle depending on the platform). Consequently, on average 35% of the images contain missing data due to cloud contamination (Roy et al., 2008). Moreover, in many parts of the world cloud contamination is much higher and temporally uneven. Under these circumstances a single cloud free image cannot be obtained for weeks or even months limiting the use of Landsat data (or any other optical low temporal resolution sensor) in many land surface parameter retrieval applications. To overcome this, methods for the multi-temporal fusion, or spatial downscaling, rely on fusing images of different sensors to overcome individual limitations by combining their properties (Roy et al., 2008; Gao et al., 2006; Gevaert and García-Haro, 2015; Inglada et al., 2016; Amorós-López et al., 2013; Liu et al., 2019; Shen et al., 2019). The Moderate Resolution Imaging Spectroradiometer (MODIS) instrument aboard the NASA EOS Terra and Aqua platforms

provides a daily revisit cycle at a coarse spatial resolution (500 m). These data are extensively used along with the Landsat missions due to their consistency (Masek et al., 2006) and their complementary specifications (Gao et al., 2006; Hilker et al., 2009a; Zhu et al., 2010; Gevaert and García-Haro, 2015; He et al., 2018), providing fine-scale spatial resolution products with better temporal sampling spanning a time series of almost twenty years.

While most of the fusion algorithms available in the literature blend vegetation indices or biophysical parameters from different sources (Hilker et al., 2009b), few methods provide the full set of calibrated radiance or reflectance estimates as outcomes. Among these, the spatial and temporal adaptive reflectance fusion model (STARFM) proposed by Gao et al. (2006) is probably the most widely used fusion algorithm. STARFM combines MODIS and Landsat images to generate daily synthetic Landsat reflectance estimates at 30 m spatial resolution, it utilizes information from neighboring pixels, which are spectrally similar to predict the missing data. Finding pixel candidates which are spectrally similar greatly increases computational costs, specially for moderate to large geographic distances (750 m) (Gevaert and García-Haro, 2015). Following a similar methodological scheme, different models that improve some aspects of STARFM have been developed. The spatial and temporal adaptive algorithm for mapping reflectance change (STAARCH) aims to better identify highly detailed spatio-temporal patterns related to land cover changes and disturbances (Hilker et al., 2009a). Zhu et al. (2010) developed an enhanced spatial and temporal adaptive reflectance fusion model to take into account differences between homogeneous and heterogeneous pixels in the prediction. However, these approaches are limited in that they rely on the availability of simultaneous cloud free Landsat and MODIS observations. As this is often unlikely, the closest gap free MODIS and Landsat images are often many days apart in the time domain (Luo et al., 2018). Another type of approach for data fusion relies on unmixing methods to infer end members and group abundances within pixels (Zurita-Milla et al., 2011; Amorós-López et al., 2013; Gevaert and García-Haro, 2015). The main advantage of this, is they do not require high and medium resolution data simultaneously for similar spectral bands. Relaxing this constraint allows for improving the spectral resolution of high spatial resolution sensors and enables the inclusion of other ancillary data sources, such as land cover data, to complement or replace high-resolution imagery in the preliminary grouping step (Zurita-Milla et al., 2011). Based on sparse representation theory, Huang and Song (2012); Song and Huang (2012) proposed a new spatio-temporal data fusion method using dictionary-pair learning based approach. The algorithms require building a correspondence between Landsat and MODIS images through learning a dictionary pair resulting in enhancing the spatial resolution of MODIS images to the Landsat spatial resolution. Other approaches for data fusion rely on sophisticated machine learning methods such as convolutional neural networks (CNN), Hopfield neural networks (HNN) and random forests (RF) (Song et al., 2018; Fung et al., 2019; Ke et al., 2016) to find the relationship between fine and coarse satellite images.

STARFM-like learning and unmixing methods for data fusion provide good accuracy, but require computationally costly spatial operations, training complex models, or unsupervised classifications as preliminary steps in their predictions. This imposes an substantial constraint, especially when applied at broad regional or global scales or over long temporal periods. As a consequence, the application of these methods is usually restricted to a reduced number of scenes covering small study areas and short periods of time (Chen et al., 2015). Alternatively, simpler and faster methods have been introduced to enable data fusion to higher spatial scales at the cost of a lower accuracy. A rule-based piece-wise regression model was proposed by Gu and Wylie (2015) to fuse MODIS and Landsat data to derive NDVI (Normalized Difference Vegetation Index) images. Simple linear models between NDVI and FAPAR (fraction of absorbed photosynthetically active radiation) were also implemented to downscale MODIS FAPAR data to

the native Landsat TM 30-m resolution using NDVI data. More recently, [Sedano et al. \(2014\)](#) proposed a pixel-based data assimilation method with a Kalman filter (KF) algorithm to fuse NDVI time series of Landsat and MODIS images. The method provided good results and accounted for the uncertainties in its calculations. Because the KF method does not require explicit parameter tuning, it is well-suited to larger scale applications. Moreover, similar KF schemes have been successfully applied to other sensors at regional scales ([Kempeneers et al., 2016](#)), as well as for the retrieval of different information such as the surface bidirectional reflectance distribution function (BRDF) and albedo ([Samain et al., 2008](#)).

Recently a variety of tools have been developed to enable large-scale processing abundant archives of publically available geospatial data. These tools enable global-scale analysis applications, such as the identification of global-scale forest cover change between 2000 and 2012 using more than 600,000 Landsat scenes ([Hansen et al., 2013](#)). Among them, the Google Earth Engine (GEE) has experienced the most rapid growth in terms of number of users and published literature ([Kumar and Mutanga, 2018](#)). GEE is in use across a wide variety of disciplines and provides user-friendly access to global time-series of satellite and vector data, cloud computing, and algorithms for processing EO data in an efficient and transparent way ([Gorelick et al., 2017](#)). Here we use the GEE cloud computing platform to implement and validate a multi-sensor data fusion approach to generate reduced noise and gap-free estimates of Landsat reflectance values. We capitalize on both the capacity of GEE to efficiently process massive amounts of data from different remote sensing data sources (MODIS and Landsat), and on an optimized bias-aware KF method in which information from observing systems and models is combined optimally to minimize residuals. The proposed Highly Scalable Temporal Adaptive Reflectance Fusion Model (HISTARFM) is optimized for land vegetation monitoring and works on a per-pixel basis. In addition, as other KF implementations, HISTARFM has the benefit of not requiring any specific parameter tuning to achieve optimal results.

In this paper, we describe the HISTARFM algorithm and we demonstrate its applicability for gap-filling and fusing the surface reflectance data from MODIS and Landsat sensors at a continental scale (contiguous United States, CONUS). More precisely, we generate gap-free monthly reflectance products at 30 m spatial resolution for six Landsat spectral bands. Finally, we perform a validation of the method for the study area and compare the results with other methods available in the literature.

2. Data

2.1. Landsat data

The Landsat missions have uninterruptedly monitored the global land surface since the launch of the first Landsat platform in 1972, constituting the longest available high spatial resolution global land surface dataset registered from space satellites ([Loveland and Irons, 2016](#)). In this work, we selected data from Landsat satellites which were concurrent with the NASA's MODIS mission to combine their information. The main characteristics of the optical bands of these Landsat sensors are shown in [Table 1](#).

The set of Landsat surface reflectance (SR) data products are routinely generated using the Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) algorithm and provided by the U.S. Geological Survey (USGS). This method was originally developed by National Aeronautics and Space Administration (NASA) and the University of Maryland ([Masek et al., 2006](#)) to generate automatically SR products. SRs are estimated for the Landsat Thematic Mapper (TM), Enhanced Thematic Mapper Plus (ETM+) and Operational Land Imager (OLI) sensors using a MODIS/6S like methodology ([Vermote et al., 1997](#)). The 6S atmospheric correction algorithm corrects for water vapor, aerosol optical thickness, ozone, geopotential height, and digital

Table 1

General description of the Landsat sensors and platforms considered in this work. Only the optical bands are shown. The bands that MODIS and Landsat sensors have in common are in bold letters.

Mission	Instrument	Time span	Bands	WL(μm)	Res(m)
Landsat 5	Thematic Mapper (TM)	1984–2013	1	0.45–0.52	30
			2	0.52–0.60	30
			3	0.63–0.69	30
			4	0.76–0.90	30
			5	1.55–1.75	30
			7	2.08–2.35	30
			8	0.52–0.90	15
Landsat 7	Enhanced Thematic Mapper Plus (ETM+)	1999–present	1	0.45–0.52	30
			2	0.52–0.60	30
			3	0.63–0.69	30
			4	0.77–0.90	30
			5	1.55–1.75	30
			7	2.09–2.35	30
			8	0.52–0.90	15
Landsat 8	Operational Land Imager (OLI)	2013–present	1	0.44–0.45	30
			2	0.45–0.51	30
			3	0.53–0.59	30
			4	0.64–0.67	30
			5	0.85–0.88	30
			6	1.56–1.65	30
			7	2.10–2.30	30
			8	0.50–0.68	15
			9	1.36–1.38	30

elevation, and provides masks for clouds, cloud shadows, adjacent clouds, land, and water. In this work, we have used the LEDAPS corrected Landsat SRs ingested in GEE, they are the highest level of image processing available for these sensors up to date, and provide ancillary data quality information and masks which help to prevent and identify data artifacts.

While Landsat 5 (ETM) and 7 (ETM+) consistency checks of their SRs have revealed, in general, a good match among them ([Claverie et al., 2015](#)), the OLI (Landsat 8) sensor shows significant differences because of adoption of more modern technology in its development. Thus, the combined use of Landsat 5, 7, and 8 data requires further adjustments to be used synergistically. We accounted for sensor differences by adjusting ETM (Landsat 5) and ETM+ (Landsat 7) bands to match OLI (Landsat 8) bands using a linear transformation as proposed by [Roy et al. \(2016a\)](#). Since the Landsat sensors always acquire their images close to nadir ($\pm 7.5^\circ$), the small variations in the bidirectional reflectance distribution function (BRDF) from changing view angles are usually not corrected ([Masek et al., 2006](#)). Following this criteria, no additional angular correction method has been applied to the Landsat data in this work. Landsat data were aggregated (mean value) to a monthly temporal resolution when multiple images were available for a given month.

2.2. MODIS data

The MODIS (Moderate Resolution Imaging Spectroradiometer) instrument is on board of the Terra and Aqua platforms. The Terra satellite was launched in 1999 and provides daily passes in the morning, while the satellite Aqua passes in the afternoon. In [Table 2](#), we show the main specifications of the used MODIS data.

As with the Landsat data, we have also used SR information. The MOD09A1 and MYD09A1 collection 6 products refers to 8-day SR composites calculated with data from the Terra and Aqua platforms respectively. Both have been corrected for atmospheric conditions using the Multi-Angle Implementation of Atmospheric Correction (MAIAC) algorithm ([Lyapustin et al., 2018](#)). The algorithm simultaneously retrieves the atmospheric variables and bidirectional reflectance parameters directly from MODIS data. Ancillary masks and quality information are provided along with the land SR bands. This information were used to discard low quality data, cloud contaminated pixels, and

Table 2

General description of the MODIS sensors and platforms considered in this work. Only the optical bands are shown. The bands that MODIS and Landsat sensors have in common are in bold letters.

Platform	Instrument	Time span	Bands	WL(μm)	Res(m)
Terra	MODIS	2000-present	1	0.46–0.48	500
			2	0.55–0.57	500
			3	0.62–0.67	500
			4	0.84–0.88	500
			5	1.23–1.25	500
			6	1.63–1.65	500
			7	2.11–2.16	500
Aqua	MODIS	2002-present	1	0.46–0.48	500
			2	0.55–0.57	500
			3	0.62–0.67	500
			4	0.84–0.88	500
			5	1.23–1.25	500
			6	1.63–1.65	500
			7	2.11–2.16	500

shadows in our blending algorithm. The MAIAC algorithm combined with the high temporal resolution of the Terra and Aqua satellites maximizes the amount of available cloud free data. MODIS data were also aggregated (mean value) to a monthly temporal resolution when multiple images were available for a given month.

2.3. Climatology and MODIS/Landsat blending

In addition to the MODIS and Landsat SR datasets, a median climatology of Landsat data and simple linear blending between Landsat and MODIS were created. These datasets were created to provide estimates when missing data is present in the scenes.

- A median monthly climatology for each considered Landsat bands were computed using the previous 10 years of data. This means, for example, that in order to produce gap filled data with HISTARFM for 2000 we will use Landsat data spanning from 1990 to 1999 to compute the climatology. The median monthly spectra were calculated by taking the median value for each month and using only gap-free high quality data according to the quality assessment (QA) information available for each Landsat scene and provided by the USGS along with the surface reflectance. Computing the median values helps to compensate extreme events within the climatology record. When no climatology could be calculated over the considered time span (i.e., permanent ice/snow or heavily clouded areas), the pixel composite value for the climatology is flagged as missing. The monthly standard deviation for each date was also computed as a first insight to the reliability of the climatology as a gap-filling approach.
- Using monthly Landsat and MODIS Terra and Aqua platforms simultaneous reflectances, a simple linear model for each band was built to relate the SRs from both sensors for the selected year of computation. The parameters of the linear model are computed in a yearly basis to assure their sensitivity to inter annual variability. The variance of the yearly linear model is used as an indicator of the quality of their estimates to predict the missing data. Further details about this simple disaggregation approach are provided in Section 3.2.1.

The LEDAPS cloud masking algorithm has been largely validated and presents good overall performance, but users have also acknowledged several weaknesses that affect any further uses of the data (Foga et al., 2017). In order to reduce the negative impacts of masking errors in HISTARFM specially over bare land and vegetated pixels, we prescribed a set of simple thresholds. Thus, pixels with negative values of Normalized Difference Vegetation Index (NDVI) (lower than -0.1),

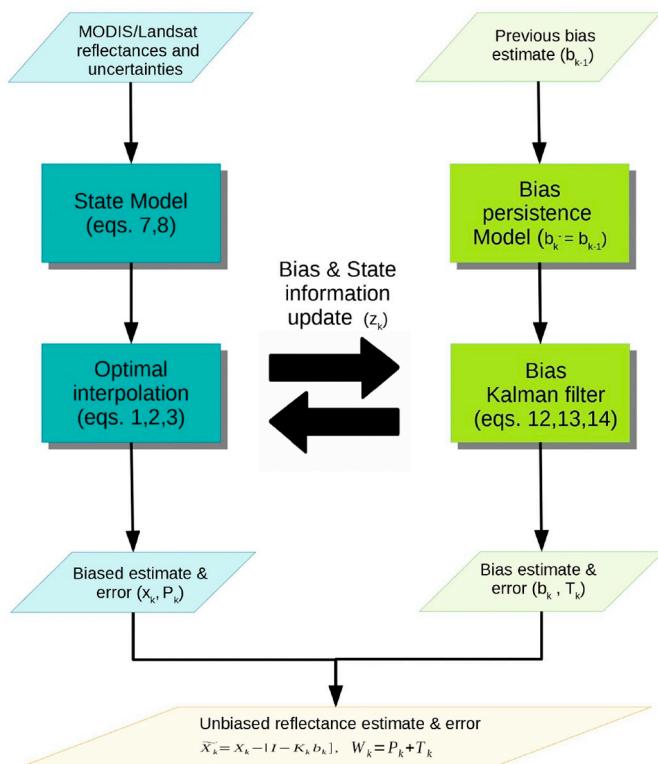


Fig. 1. Flowchart illustrating the data assimilation approach presented in this work (HISTARFM).

surface reflectance values higher than 0.4 in the visible bands, and values higher than 0.7 for NIR and IR bands were masked. These thresholds were empirically introduced to discard misidentified snow/ice and clouds by the LEDAPS algorithm but, in some cases, they could erroneously mask very bright soils and artificial areas.

3. Highly scalable temporal adaptive reflectance fusion model (HISTARFM)

The HISTARFM algorithm relies on a bias-aware Bayesian data assimilation scheme implemented in GEE. The proposed approach uses two linked estimators operating synergistically to filter out random noise and reduce the bias of Landsat spectral reflectances (see Fig. 1). The first estimator is an optimal interpolator (shown in blue color in the flowchart and explained in detail in Section 3.2) that produces estimates of Landsat reflectance values for a given time by combining Landsat climatology, pre-computed from the available Landsat record, and fused MODIS and Landsat reflectances obtained from overpasses closest to the time of interest. The fusion of reflectances is achieved using an efficient pixel-wise linear regression model. The second linked estimator is a Kalman filter (shown in green color in the flowchart and explained in detail in Section 3.3) that corrects the bias of the reflectance produced by the first estimator.

3.1. Kalman filter and Bayesian estimation

Errors in a time series of remotely sensed reflectances can be decomposed into the sum of a random component with zero mean and no temporal correlation (zero-mean i.i.d. process), and a systematic component (bias) that may vary in time and offsets the error mean away from zero. Existence of random and systematic errors is often the case in optical remote sensing data due to unfavorable atmospheric conditions that generally decrease near-infrared reflectance and increase reflectance in visible spectral bands (Goward et al., 1991; Kobayashi and

Dye, 2005). This undesired noise propagates to higher level products based on these spectral bands. For instance, it can lead to spurious and unrealistic fluctuations in commonly used indices that use ratio of optical bands, such as the NDVI.

An effective way to ameliorate this problem is to reduce the reflectance noise using a KF or other standard statistical estimators (Chen et al., 2004; Julien and Sobrino, 2010; Moreno et al., 2014). However, Kalman filters and standard Bayesian estimators assume that observation and process errors are Gaussian and unbiased (i.e. zero mean errors). If the process forecasts or the observations are biased, the correction process will not be optimal, the posterior estimate of the mean may not be accurate, and the posterior variance may be incorrectly estimated because the posterior mean and covariance are conditioned on process forecasts and on observations. Specifically, the presence of forecast and observation biases violate the assumptions that guarantee the optimality of the filter and can degrade its performance and even cause filter divergence in the case of large bias-to-variance ratios. For this, our data assimilation methodology increases filter accuracy by linking two estimators that recursively reduce the errors and bias of reflectance estimates, as it will be further explained in Section 3.3. To facilitate the reading of the methods section a table of symbols can be found in Appendix A.

3.2. Step one: Optimal interpolation

We use a standard Bayesian estimator to blend an a priori estimate of reflectance with Landsat observations of a given month k :

$$\mathbf{K}_k = \mathbf{P}_k^{-} \mathbf{H}^T (\mathbf{H} \mathbf{P}_k^{-} \mathbf{H}^T + \mathbf{R})^{-1} \quad (1)$$

$$\mathbf{x}_k = \mathbf{x}_k^{-} + \mathbf{K}_k (\mathbf{z}_k - \mathbf{H} \mathbf{x}_k^{-}) \quad (2)$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{H}) \mathbf{P}_k^{-}, \quad (3)$$

where the subscript k denotes dynamic variables at the k time step (monthly), \mathbf{x}_k^{-} is the a priori estimate of reflectance, \mathbf{z} is the Landsat observation of reflectance, \mathbf{K} is the Kalman gain which represents the relative weight given to the observations (\mathbf{z}) and the a priori estimate (\mathbf{x}_k^{-}), \mathbf{P}_k^{-} is the error covariance of the prior estimate, \mathbf{H} is the observation operator or measuring matrix that describes how model outputs relate to observations, \mathbf{R} is the Landsat error covariance, \mathbf{x} is the corrected reflectance (posterior estimate), and \mathbf{P} is the error covariance of the posterior estimate. Since we assume that all variables are normally distributed, Eqs. (1)–(3) form a static Bayesian estimator that differs from a standard Kalman filter in that there is no dynamic model that produces a priori estimates of reflectance. A dynamic model is not needed because the priors are given at time k by an external source, as we describe below. Without the dynamic model, these equations are simply the update equations of the standard KF.

3.2.1. Production of the a priori estimate \mathbf{x}_k^{-}

Before Landsat observations are available at month k for assimilation, the best estimate of reflectance at that month is given by our prior \mathbf{x}_k^{-} . The prior, therefore, must be available even when Landsat observations at month k do not exist. We generate this a priori estimate of reflectance by combining the climatology of Landsat reflectances (mean and variance of the 10 years preceding month k) with reflectances from the MODIS sensor using a Bayesian Model Averaging (BMA) approach.

Before MODIS reflectances can be combined with the Landsat climatology, they need to be downsampled to the spatial resolution of Landsat. The spatial disaggregation of MODIS reflectance is achieved using a linear mapping model with coefficients β_i determined using the ordinary least squares solution of a pixel-wise MODIS-to-Landsat linear regression model between all available MODIS and Landsat images for the selected year of computation:

$$\beta_i = (\mathbf{U}_{i,MOD}^T \mathbf{U}_{i,MOD})^{-1} \mathbf{U}_{i,MOD}^T \mathbf{z}_i, \quad (4)$$

where $\mathbf{U}_{i,MOD}$ is the augmented input matrix $[\mathbf{1}, \mathbf{u}_{i,MOD}]$, $\mathbf{1}$ is an all-ones vector, $\mathbf{u}_{i,MOD}$ is a vector of the gap free MODIS reflectances for the 12 months in the selected year of computation and the 30 m Landsat pixel i , \mathbf{z}_i is a vector of the corresponding Landsat reflectances for the same 12 months, and vector β_i contains the pair of regression coefficients (intercept and slope) specific for each pixel i and year of computation. MODIS reflectances are resampled at 30 m using a nearest neighbor algorithm to obtain one-to-one pixel i between MODIS and Landsat images. This blending algorithm has been successfully used in previous work to spatially disaggregate MODIS NDVI (He et al., 2018, 2019).

Once the β_i coefficients are determined, the spatial disaggregation of MODIS at any month k is obtained applying the linear model:

$$u_{i,k,MOD}^{30} = \mathbf{U}_{i,k,MOD} \beta_i, \quad (5)$$

where $u_{i,k,MOD}^{30}$ is the MODIS reflectance at pixel i and month k spatially disaggregated to 30 m and $\mathbf{U}_{i,k,MOD} = [\mathbf{1}, u_{i,k,MOD}]$, with $u_{i,k,MOD}$ being the reflectance at pixel i and month k of the MODIS image resampled at 30 m using the nearest-neighbor method.

We assume that the linear model is unbiased (i.e. the mean of the disaggregated MODIS reflectances equal the mean of the Landsat reflectances), therefore the error covariance of the disaggregated MODIS reflectance is:

$$\mathbf{P}_{k,MOD} = \mathbb{E}[(\mathbf{z}_i - \mathbf{U}_{i,k,MOD} \beta_i)^T (\mathbf{z}_i - \mathbf{U}_{i,k,MOD} \beta_i)]. \quad (6)$$

The downsampled MODIS reflectance is subsequently blended with the long-term expectation of reflectances (Landsat climatology) using BMA. BMA accounts for model uncertainties to provide better gap-filling and predictions than an individual model (Hoeting et al., 1999). BMA uses the outputs and variances of the two sources of information to reduce the uncertainties, and provides more robust estimates of the mean and variance of surface reflectances. The prior estimates of reflectance mean and covariance used in (2,3) are finally given by:

$$\mathbf{x}_k^{-} = \bar{\mathbf{z}}_k \frac{\mathbf{P}_{k,MOD}}{\mathbf{P}_{k,LS} + \mathbf{P}_{k,MOD}} + \mathbf{u}_{k,MOD}^{30} \frac{\bar{\mathbf{P}}_{k,LS}}{\mathbf{P}_{k,LS} + \mathbf{P}_{k,MOD}}, \quad (7)$$

$$\mathbf{P}_k^{-} = (1 - \gamma) \left(\frac{1}{\bar{\mathbf{P}}_{k,LS}} + \frac{1}{\mathbf{P}_{k,MOD}} \right)^{-1}, \quad (8)$$

where $\bar{\mathbf{z}}_k$ is the Landsat climatological mean of the 10 years prior to month k , $\bar{\mathbf{P}}_{k,LS}$ is the Landsat climatological variance of the 10 years prior to month k , γ is a fraction of the error covariance of the estimate that is attributed to bias (see next section), and $\mathbf{P}_{k,MOD}$ is the variance of the downsampled MODIS reflectance, Eq. (6).

3.3. Step two: Bias correction

The second linked estimator corrects the estimation biases using a separate bias Kalman filter without feedback (Friedland, 1969; Dee and Da Silva, 1998; Driouct et al., 2006). In our implementation, the bias correction (SepKF) uses as observations the difference between observed reflectance and their estimated counterparts from the previous step. It also links the error covariance of the posterior estimate \mathbf{P} obtained in the previous step (Eq. (3)) assuming unbiasedness, the actual error covariance of the estimate, \mathbf{W}_k , and the bias error covariance matrix, \mathbf{T} :

$$\mathbf{P}_k = \mathbf{W}_k - \mathbf{T}_k, \quad (9)$$

where the state and bias estimates (steps one and two) are assumed to be uncorrelated.

Note now that the mean bias and error covariance estimates are propagated in time with a persistence model, which implies that the bias forecast at time k is the corrected (posterior) bias from time step $k - 1$. The bias forecast error covariance, \mathbf{T}_k , is considered a fraction γ of the prior reflectance estimation error as in Eq. (8):

$$\mathbf{b}_k^- = \mathbf{b}_{k-1} \quad (10)$$

$$\mathbf{T}_k^- = \gamma \left(\frac{1}{\mathbf{P}_{k,LS}} + \frac{1}{\mathbf{P}_{k,MOD}} \right)^{-1}, \quad (11)$$

where γ controls the amount of information from observations used in the bias filter. Setting $\gamma = 0$ makes SepKF equivalent to the classical KF (white Gaussian noise), while $\gamma = 1$ implies that all the error is attributed to the bias (deterministic error). Please, note that γ controls the sensitivity of the bias filter and the speed at which the filter converges to the actual bias, but does not determine the final estimate. This parameter is usually determined empirically but satisfactory results have been reported for any γ value lower than 1 (Drécourt et al., 2006). In this paper we set $\gamma = 0.6$. Our experience is that this value captures the trade-off between land cover pixels that change rapidly and tend to have highly biased reflectances, such as croplands, and pixels that have smooth and slow-varying transitions, such as the case of unmanaged forests. The bias correction equations can thus be written as:

$$\mathbf{L}_k = \mathbf{T}_k^- \mathbf{H}^T (\mathbf{H} \mathbf{T}_k^- \mathbf{H}^T + \mathbf{R})^{-1} \quad (12)$$

$$\mathbf{b}_k = \mathbf{b}_k^- - \mathbf{L}_k (\mathbf{z}_k - \mathbf{H}(\mathbf{x}_k^- - \mathbf{b}_k^-)) \quad (13)$$

$$\mathbf{T}_k = (\mathbf{I} - \mathbf{L}_k \mathbf{H}) \mathbf{T}_k^-, \quad (14)$$

and the final unbiased estimate of the mean reflectance, which combines the results of both filters is:

$$\bar{\mathbf{x}}_k = \mathbf{x}_k - (\mathbf{I} - \mathbf{K}_k \mathbf{H}) \mathbf{b}_k \quad (15)$$

where, as per Eq. (9), the covariance error of the unbiased mean estimate becomes:

$$\mathbf{W}_k = \mathbf{P}_k + \mathbf{T}_k, \quad (16)$$

4. Results

In this section we show continental scale images processed with HISTARFM at full spatial resolution to demonstrate that the algorithm runs operationally in GEE. In addition, a continental scale validation of the results has been also carried out to assess the limitations of the method. Finally, we compare results of the algorithm with one of most widely used data fusion algorithms available in the literature (STARFM) and analyze the possible sources of the observed discrepancies.

4.1. Example of application over the contiguous United States

To illustrate and validate the proposed KF method for smoothing and gap filling Landsat data, we processed a full year (2010) covering the full CONUS and exported the outcomes as a monthly images as assets in GEE. The dataset was processed at Landsat native spatial resolution (30 m) and provides the estimated reflectances along with their associated predicted uncertainties. The monthly temporal resolution was chosen to reduce data storage needs, this point is especially critical when working at continental scales as we do in the present work. In Fig. 2, we illustrate an example of the differences between the original Landsat reflectance bands (RGB) and the results of the proposed method for April 2010. It is noteworthy to mention that in general terms while the original Landsat data contain numerous gaps due to cloud contamination and sensor malfunctioning (Landsat 7), the proposed method seems to be satisfactory recovering the missing information and preserves most spatial patterns. In addition, the gap filled bands also exhibit slightly smoother spatial patterns than the original Landsat data. This is a desirable result of combining information from different sources (Landsat climatology and actual Landsat data) and sensors (MODIS), and it is generally more accurate than using a single sensor reflectance measurements themselves (Kalman, 1960). In spite of the mentioned improvements, some undesired variability is still noticeable in the processed data. This variability could be related with Landsat

view zenith BRDF effects which were not corrected and are reported to be potentially not insignificant (Ju et al., 2012), but also related with bad cloud/shadow or snow masking in any of the considered remote sensing data products included.

As mentioned above, the proposed method also provides uncertainty estimates along with the predicted reflectance as ancillary data for error propagation purposes. Error propagation methods allow to quantify the effect of variables' uncertainties on the total uncertainty of any function of the input variables. As an example of application, we computed the Normalized Vegetation Index (NDVI) and calculated its propagated uncertainty $\sigma[\text{NDVI}]$ function of the uncertainties of the red $\sigma[B_3]$ and near infrared $\sigma[B_4]$ bands, assuming that both bands are independent variables:

$$\sigma[\text{NDVI}] = \frac{2}{(B_3 + B_4)^2} \sqrt{B_4^2 \sigma[B_3]^2 + B_3^2 \sigma[B_4]^2} \quad (17)$$

In Fig. 3, we show the computed NDVI using the gap filled and smoothed reflectances. As shown in Fig. 2, the monthly composite presented significant gaps and noise in the considered area. Despite applying HISTARFM to remove noise from the signal and obtaining improved gap free results, there is still some noticeable striping in certain parts of the processed images. These artifacts are mostly related with faulty thin cloud identification in the original LEDAPS algorithm. This misidentification does negatively affect HISTARFM performance, because the algorithm considers as good data information which is not so reliable. In any case, the NDVI presents a smoother spatial variability than the RGB reflectances and also no presence of cloud related gaps. A smoother spatial variability is obtained by the NDVI as a result of its normalization effect, which helped to mitigate some of the undesired and non corrected illumination effects and bad masking. The NDVI error shows clear spatial patterns. In most cases, the highest errors are located over the highly elevated terrains and high latitudes. These areas are characterized by having high snow occurrence during the winter and early spring months, which cause a lot of undesired variability and gaps still present in the figure shown (April) due to full or partial snow contamination. Visually inspecting the predicted NDVI uncertainties with the original non gap filled data shown Fig. 2, it is observed that predicted uncertainties are higher when Landsat information is not available. This is a consequence of HISTARFM using only the estimated reflectances (MODIS and the Landsat climatology) to compute its predicted values instead of using also concurrent Landsat measurements. In particular, the latter are prioritized by the model when available, because the Kalman filter encourages the use of data which present higher accuracies to decrease the posterior variance in the final estimates (see Eqs. (1), (2), (3)).

4.2. Validation over a selection of sites

To evaluate the quality of the proposed approach at a continental scale, 1050 locations were randomly selected over the continental US. These sites were chosen to represent the most common vegetation types present in the area according to the US NLCD (National Land Cover Database) for the year 2011 (Homer et al., 2015). This land cover is provided at 30 m spatial resolution, has been optimized to be nationally consistent for the US, and was produced in 5 epochs (1992, 2001, 2006, 2011, and 2016). The sites were selected to be homogeneous in a 3 by 3 pixel grid and consist on 150 locations of each of the following land covers: deciduous forest (DF), evergreen forest (EF), mixed forest (MF), Shrub/scrub (SH), grassland/herbaceous (GR), pasture/hay (PA), and cultivated crops (CR). The artificial gaps were then generated in the 1050 Landsat time series removing 15% of the available data. On top of these artificially created gaps, we have the gaps that were originally present in the data due to natural factors such as clouds or snow, yielding a mean percentage of missing data of 23%. After creating the artificial gaps, the method was run over these time series and its estimations and predicted uncertainties were compared with the

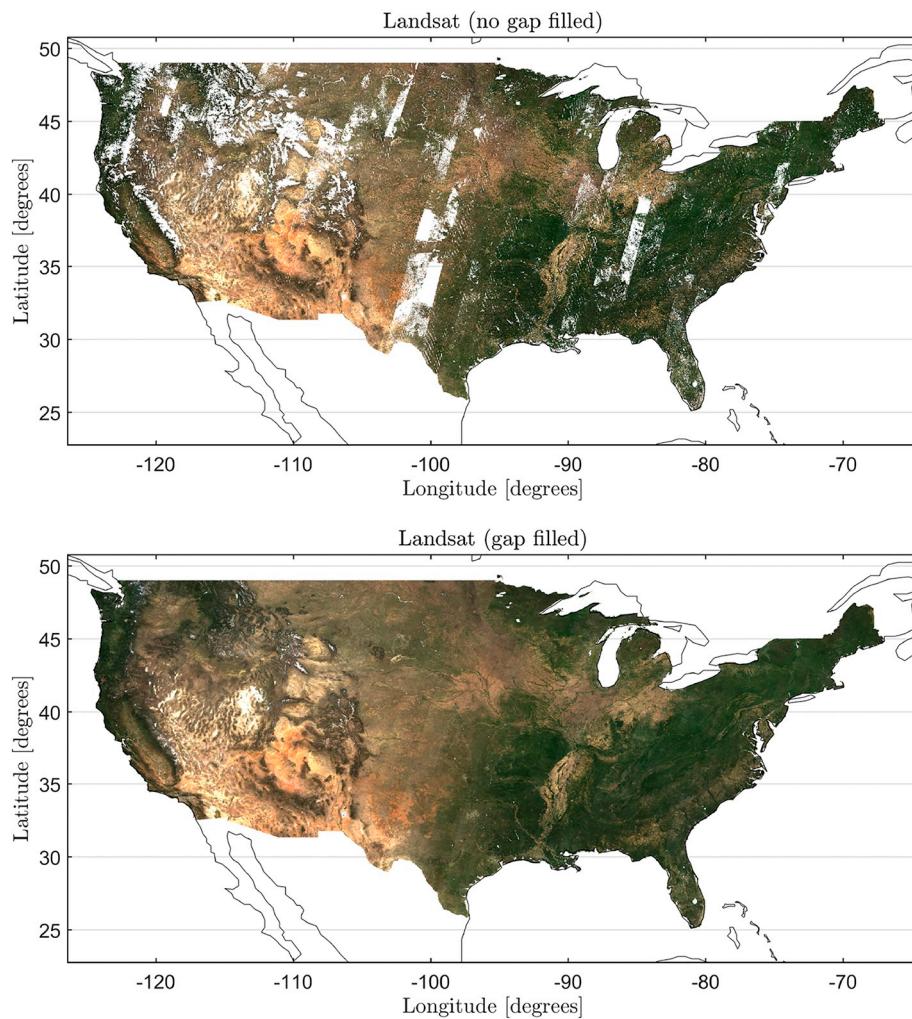


Fig. 2. RGB composites with original Landsat LEDAPS reflectance (top) and the smoothed and gap filled reflectance estimates by HISTARFM (bottom) in April 2010.

deliberately removed data for validation purposes.

Table 3 summarizes the calculated errors over the validation dataset. The relative mean errors (rME) remain below 1.5% for all the bands, indicating that our estimates are unbiased when compared with actual Landsat reflectance values. Moreover, correlations are significantly high between observed and predicted values for all considered Landsat bands, while the relative mean absolute errors (rMAE) and relative root mean squared errors (rRMSE) are low to moderate and vary significantly between bands. These findings are also consistent with previously shown results, being the magnitude and variability of rMAEs for the different bands similar to those reported in the literature for Landsat LEDAPS algorithm (Claverie et al., 2015). We therefore speculate that these uncertainties might be partly attributed to the original Landsat surface reflectance data we want to predict and not only to the proposed method. B4 (NIR) and B5 (IR) bands yielded to the lowest relative errors, while the visible bands obtained the highest errors in relative terms. This difference in performance for the visible bands is even more noticeable in the blue band (B1), and it is related with more pronounced atmospheric impacts in this part of the spectrum (Vermote and Kotchenova, 2008). Fig. 4 shows the predicted versus observed Landsat data scatter plots for each band separately. As above mentioned, our results show good correlations and virtually no bias for all the bands. The scatter plot figures illustrate that most of the values are spread around the one-to-one line (black dotted line). This corroborates visually the absence of bias in our estimates and low dispersion over the range of variability of all considered spectral bands.

The consistency between the errors over the validation data set and the predicted uncertainties have been also compared. Fig. 5 summarizes the results of the errors and the predicted uncertainties by spectral band and vegetation type. Although there is a slight underestimation of the predicted RMSE for the highest values, it is important to note that the degree of consistency between the actual RMSE (left) and the predicted RMSE (right figure) is high either for both the different bands and plant functional types. This provides the foundation for using the predicted uncertainties for error propagation purposes (see Section 4.1) in future applications of the provided gap filled reflectances. B4, B5, and B6 bands present the highest RMSE values between the predicted and reference data, while the visible bands (B1, B2, and B3) have the lowest estimated error. This difference is also accurately captured by the predicted uncertainties of the proposed approach. It is important to note that, in general, the mean reflectances for the visible bands are significantly lower (50%) than the mean reflectance values for the near and medium infrared, so, in relative terms, visible bands tend to be the most uncertain as also shown before in Table 3. Forests vegetation types present the lowest errors while grasslands, pastures and croplands have the highest RMSE values. We speculate that these differences could be attributed to two main factors. First, as previously mentioned, one of the sources we are using for predicting the missing information is a climatology of Landsat reflectances. Forests vegetation types are less prone to be affected by human type disturbances in comparison with crops which are totally man managed. This yields forest types to present less inter annual variability than cropland areas, explaining why the

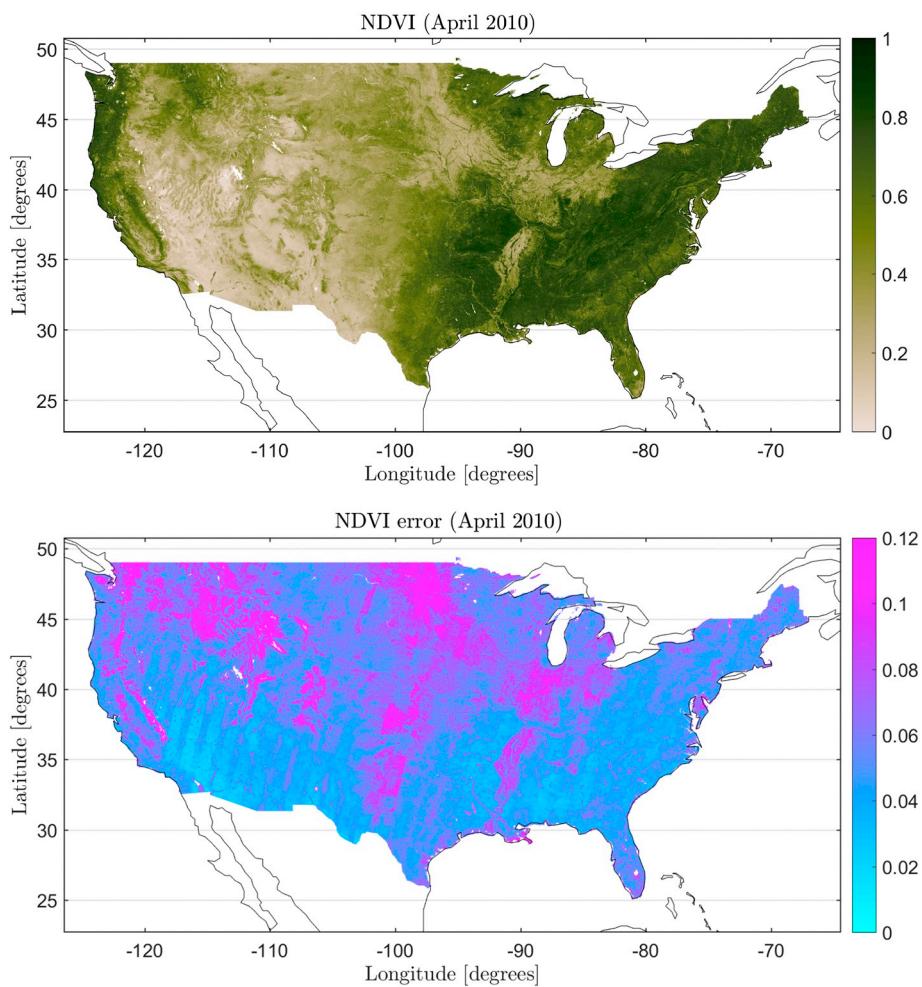


Fig. 3. Calculated NDVI values with the gap filled data and their associated uncertainties. The ranges of NDVI and the NDVI uncertainties have been constrained for illustration purposes.

Table 3

Summary of the results over the validation data set. Relative values are in %.

Band	ME	MAE	RMSE	rME	rMAE	rRMSE	R
B1	0.0009	0.011	0.017	1.5	17	29	0.85
B2	0.0003	0.011	0.018	0.5	13	22	0.90
B3	0.0002	0.015	0.023	0.16	16	25	0.92
B4	0.0028	0.026	0.039	1.1	10	16	0.87
B5	-0.0004	0.024	0.037	-0.16	10	16	0.91
B7	-0.0006	0.022	0.035	-0.4	15	23	0.91

computed climatology could be a more suited predictor for forests than for crops. Second, grasslands, pastures, and crops have a significantly faster growth than forest types, yielding to higher variance at the considered temporal resolution (monthly time steps). This increase in variance is, in fact, not attributed to noise exclusively and it could be actually related with fast changes in vegetation canopy which occur within a monthly time step.

Even though we have used the ancillary Landsat quality band to mask and discard cloud or snow pixels over land to create the validation dataset and to run HISTARFM. Snow, cloud, and cloud shadow detection in satellite imagery is a complicated task and, very often, significant cloud and snow contamination is still present in data after carrying out a masking processing step. The Landsat atmospheric correction algorithm (LEDAPS) is not an exception to these problems, and usually faulty cloud and snow screening are spectrally characterized by having high visible (B1, B2, and B3) and NIR bands (B4) reflectance

values compared to soil and vegetation canopy spectral response for which shortwave bands (B5 and B7) present more similar values. Moreover, if very dense aerosols and haze/thin clouds are present the scattering effects may also contaminate significantly B4, B5, and B7 bands (Liang, 2005). The analysis of the temporal distribution of the errors for all considered bands is shown in Fig. 6. These results in concordance with the above mentioned masking problems, and thus, for example, the visible bands have the highest relative errors and, as expected, these errors increase during winter and spring months when generally there is more snow and cloud occurrence in the study area (Roy et al., 2016a). Furthermore, as also reported in the literature, the blue band (B1) has the highest errors as result of its extra sensitivity to this specific kind of contamination.

As mentioned in Section 3, the implemented Kalman filter approach assumes that observation and process errors are Gaussian and unbiased. We examined our residuals to assess if these assumptions were fulfilled to guarantee that the choice of model was appropriate. Fig. 7 shows the normal probability plots of the residuals over the created validation dataset. In normal probability plots if the residuals belong to a normal distribution the data points have to be spread out along the reference line (red line in our case). On the contrary, any other kind of distribution other than normal introduces curvature in the data plot creating a departure from the indicated reference line. In our case, the residuals of all the bands belong very likely (between the 10th and 90th percentiles) to a normal distribution, but the normality of the residuals breaks at the extremes for the least probable cases. Moreover, the residuals of the visible bands tend to be more normal than the NIR and

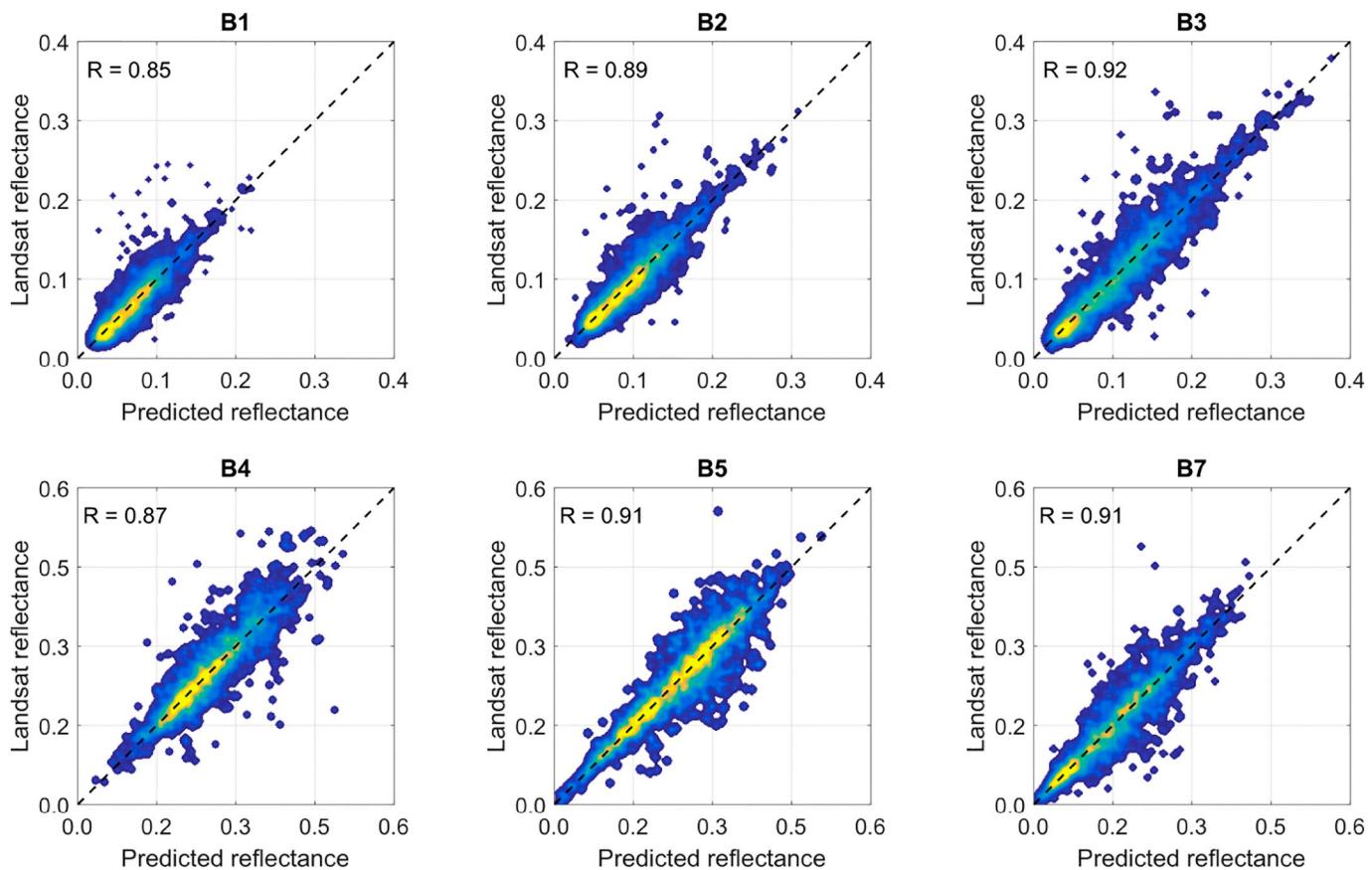


Fig. 4. Scatter plots of the predicted versus observed Landsat reflectances. The one-to-one line (black) is shown for reference.

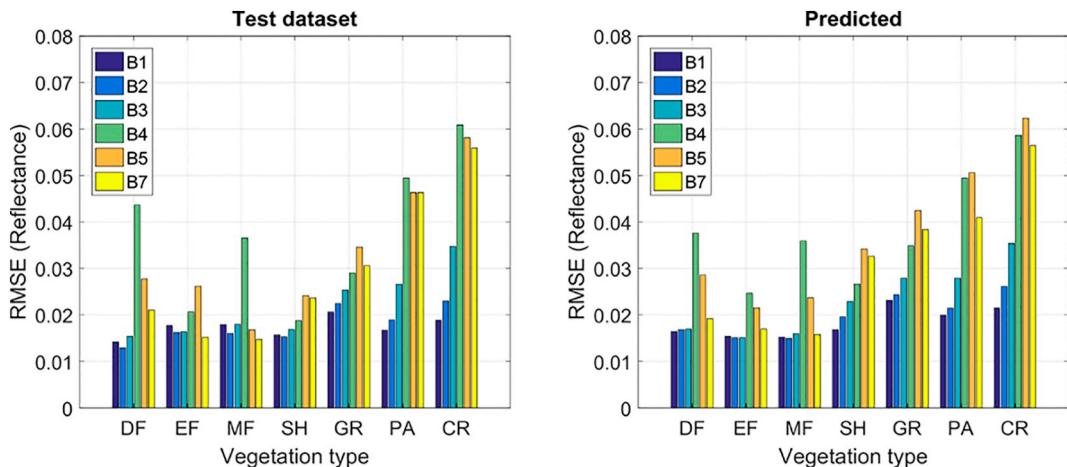


Fig. 5. RMSE values over the validation data set (left) and the RMSE estimated by the proposed method (right). The considered vegetation types are: deciduous forest (DF), evergreen forest (EF), mixed forest (MF), Shrub/scrub (SH), grassland/herbaceous (GR), pasture/hay (PA), and cultivated crops (CR).

MIR bands. So, even though the carried out analysis indicated that model residuals are not fully Gaussian, their normality is preserved in most cases indicating that the followed approach is still suitable for the proposed problem. In addition, because the lack of normality appears at high absolute values in the residuals and the predicted uncertainties have been proved to be reliable, the predicted uncertainties could be also used to discard bad quality data in further applications.

4.3. Comparison with STARFM

With more than 700 cites (<https://www.scopus.com/>), the

STARFM algorithm is probably the most widely used data fusion methods available in the literature. The source code, documentation, and executable files to run STARFM can be downloaded from <https://www.ars.usda.gov/>. With the objective of comparing how our method and STARFM algorithms perform under different circumstances, we processed the visible (B1, B2, and B3) and near infrared bands (B4) which are most commonly used in many applications, and selected five locations with different climatic, vegetation, and topographical characteristics in the study area (CONUS). The first site (zone 1) corresponds with a cropland area located in Moses Lake city, Washington (US). The second site (zone 2) corresponds with a high

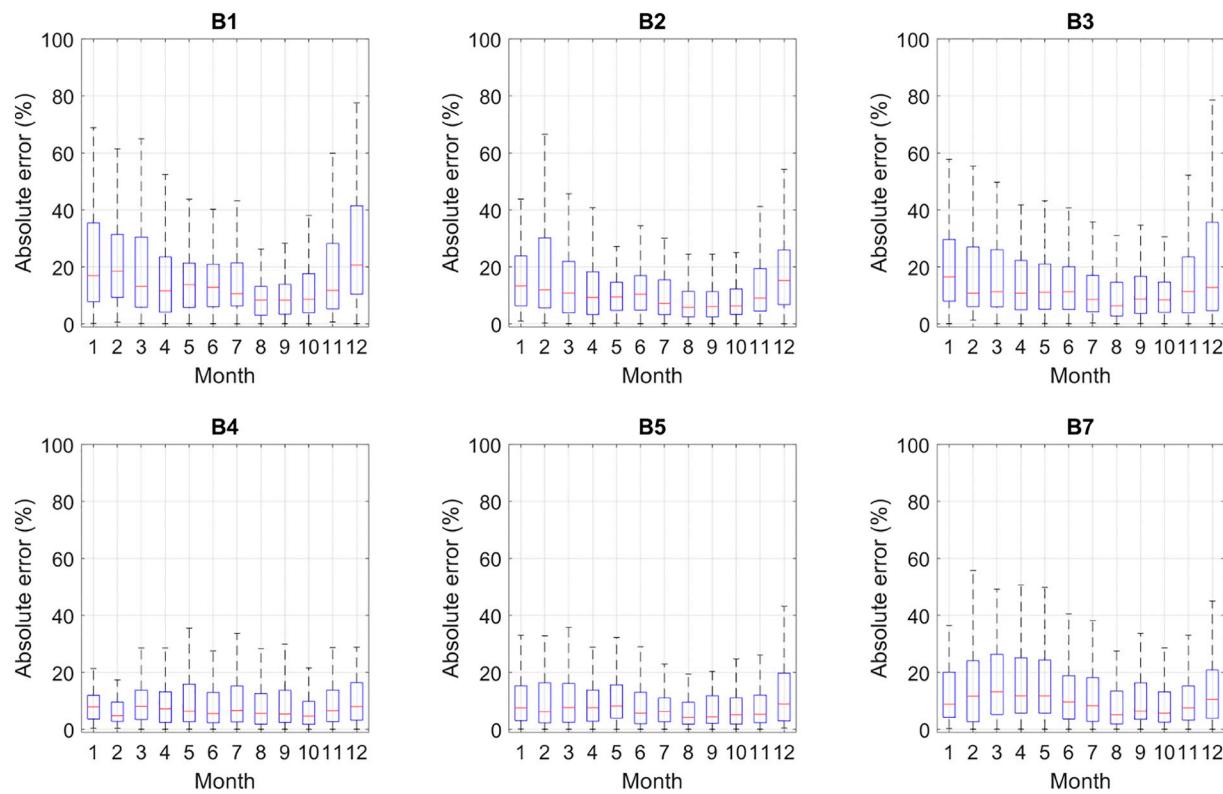


Fig. 6. Box plots showing the temporal evolution of the relative absolute errors for the different Landsat bands.

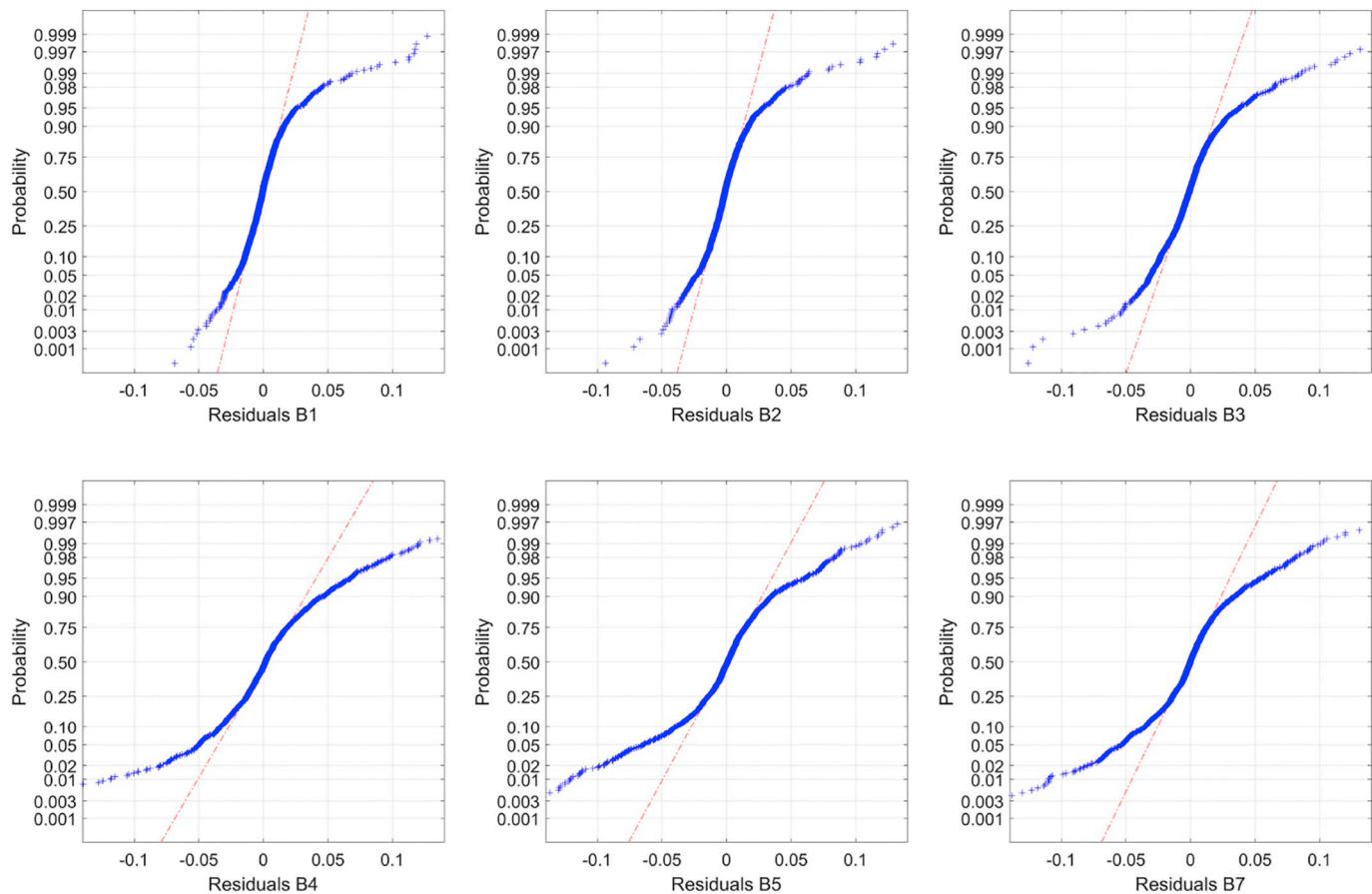


Fig. 7. Normality tests of the residuals for the considered Landsat bands.

elevation (1500 m) and complex mountainous area in Selkirk, Washington state (US), which is mostly dominated by evergreen forest vegetation. The third site (zone 3) is located in Brush Creek Township, Ohio state (US), where this area is mainly dominated by deciduous forest vegetation. The fourth considered site (zone 4) is located in Washington D. C., District of Columbia (US), the area is a very heterogeneous urban area which is mostly dominated by a mixture of constructed materials and vegetation, in addition, the site includes big patches of deciduous forest dominated by big trees with a strong seasonal variation. Finally, zone 5 is located close to Tonopah town, Nevada (US), this area has an arid, cold desert climate with cool winters and hot summers, it is mostly and equally covered by two vegetation types: shrubs less than 5 m tall and not densely populated evergreen forests. For the chosen locations we extracted three monthly composites of Landsat and MODIS multi-band images corresponding with the months July, August and September of 2018.

In order to be able to quantitatively compare both algorithms, we gap filled full images of Landsat data for the selected sites. This way we had large spatial gaps of ground truth to assess the gap filling done by the algorithms. STARFM was run in a local computer using the July and September months to predict the intermediate month (August) with the standard configuration of the parameters proposed by the authors in the provided documentation with STARFM code. Additionally, we ran HISTARFM in GEE. The predictions of our method and STARFM were then compared against the available cloud and snow free Landsat monthly data over the five sites to qualitatively and quantitatively evaluate both data fusion algorithms under the same circumstances.

Fig. 8 shows false color RGB images (Landsat bands B1, B2, and B3) obtained with the original Landsat data to be gap filled, the estimations of STARFM, and the estimations of the proposed algorithm. Our method compares well with STARFM and, most importantly, with the original Landsat data. Moreover, HISTARFM seems to be more robust to noise than STARFM and presents, generally, smoother spatial variability. This effect is even more remarkable in certain areas of the images, such as the red artifacts in STARFM predictions in the center of cropland site (zone 1), the dark areas all over the zone 3, or the stripping effect due to the faulty Landsat 7 sensor also more noticeable in the zone 5. It is also worth to mention that the original zone 2 Landsat image looks significantly less contrasted than the predictions of both methods. We hypothesize it could be attributed to contaminated Landsat radiances due to thin clouds which were not correctly identified by the LEDAPS algorithm. Moreover, a significant discontinuity between acquisitions is easily noticeable between adjacent Landsat scenes in the lower right part of the image. This sudden change is not observed in previous or posterior satellite acquisitions, and therefore can not be attributed with natural changes of the vegetation within the considered month.

Table 4 shows the results of the comparison of STARFM and HISTARFM estimates for August 2018 in the five considered study areas, we used the available Landsat data for that date to estimate the accuracy of both methods recovering missing data. Generally speaking, our approach outperforms STARFM in any of the considered zones and analyzed spectral bands as can be seen just only for the number of emphasized statistics in the table. As above mentioned, the qualitative inspection of the results in **Fig. 8** indicated that the HISTARFM algorithm was able to capture better the spatial variability in all the locations. **Table 4** numerically corroborates this and, most importantly, correlations are equal or significantly higher than those obtained by STARFM in all analyzed bands and zones. In addition, RMSE values are the lowest for our approach in the majority of cases and close in magnitude to their MAE values, suggesting that HISTARFM estimations are less noisy than the STARFM ones. On the contrary, the analysis of the bias in the residuals doesn't indicate a clear winner, obtaining both methods low biased results independently of the band or the location under study. The highest errors and lowest correlations for both gap filling methods simultaneously are obtained for the zone 3. Finally, the larger discrepancies between HISTARFM and STARFM are observed in

zone 4 and it could be attributed to the fact that STARFM depends on temporal information from pure homogeneous patches of land cover at the MODIS pixel scale. Zone 4 is the most heterogeneous site and, under these circumstances, it has been documented that STARFM predicted results can be misleading (Gao et al., 2006; Hilker et al., 2009a)

In addition to the validation statistics computed for each study area, we also compared how the different methods affected the distribution of reflectance values. **Fig. 9** shows the normalized histograms for the bands most frequently used in vegetation monitoring (B3 and B4). As can be seen, the differences among distributions vary significantly depending on the band and the characteristics of the study area. The zone 1, which is a very heterogeneous site, shows bigger differences between STARFM and HISTARFM in the red band (B3) than in the near infra red (B4). This could be related with a higher sensitivity to noise of STARFM, as can be seen in unrealistically reddish artifacts we pointed out in **Fig. 8**. Zone 2 and zone 5 distributions present the highest consistency between the two bands in both methods. On the other hand, zone 3 and zone 4 show significant discrepancies in both considered bands, being the deviation in the mean values of the distributions in band B4 more than 20% for the zone 3.

With the aim of visualizing the impact of the discrepancies in the predicted reflectances (**Fig. 9**) for land vegetation monitoring, we computed the NDVI with the available Landsat data in August 2018, the bands predicted with STARFM, and the bands predicted with HISTARFM. **Fig. 10** shows the distributions of the NDVI residuals for both approaches. As expected from previous analyses, HISTARFM outperforms STARFM in all five sites for the bands shown. The distribution of the residuals resulted in more narrow Gaussian shaped distributions which are also closer to zero deviation in general. As show in **Table 3**, in zones 1 and 5 both approaches obtained the best results and rather low NDVI differences. The residuals in zone 2 are very similar for both methods, but high biases in the estimated NDVIs are observed. This could be attributed to cloud contamination which affects the NDVI signal creating unrealistic sudden drops in the time series. The comparison of the differences between the distributions of the reflectances (**Fig. 9**) and the residuals of the NDVI in zone 3 (**Fig. 10**) suggests a considerable improvement of our approach when compared with STARFM, being the mean of the residuals lower and with less variability. Finally, for the zones 4 and 5, STARFM presents slightly lower mean NDVI biases than HISTARFM but higher variance in its estimates, this is also shown for the spectral bands used to compute the NDVI in **Table 3**.

Finally, we also analyzed the temporal evolution of the errors for HISTARFM and STARFM. For two of the selected locations (zones 4 and 5), we extracted 6 monthly composites of Landsat and MODIS multi-band images corresponding with the months between March and September of 2018. In order to be able to quantitatively compare both algorithms, we gap filled full images of Landsat data for the selected sites and for the months May, June, July and August. We ran both HISTARFM and STARFM removing these specific dates and compared them with actual Landsat observations. **Fig. 11** shows the RMSE for each band and four consecutive months when temporal variability due to phenology changes is high. These results highlight the better performance of HISTARFM against STARFM for all considered dates and bands in both zones, but also show the high sensitivity of STARFM to noise in its inputs due to faulty cloud masking or aerosol contamination and the robustness of HISTARFM to it. For the two considered zones, STARFM presented massive drops in performance during May but also in June in zone 4. A closer inspection of the input data used to run STARFM predictions revealed the presence of thin clouds contamination and significant cloud gaps. It seems that missing and faulty information degrades STARFM's performance more drastically than in HISTARFM. This is an expected result, since STARFM heavily relies on two observations to predict missing data. Moreover, STARFM's high sensitivity to the quality of its inputs is also an important constraint that has implications for its practical use, as reliable predictions can only be

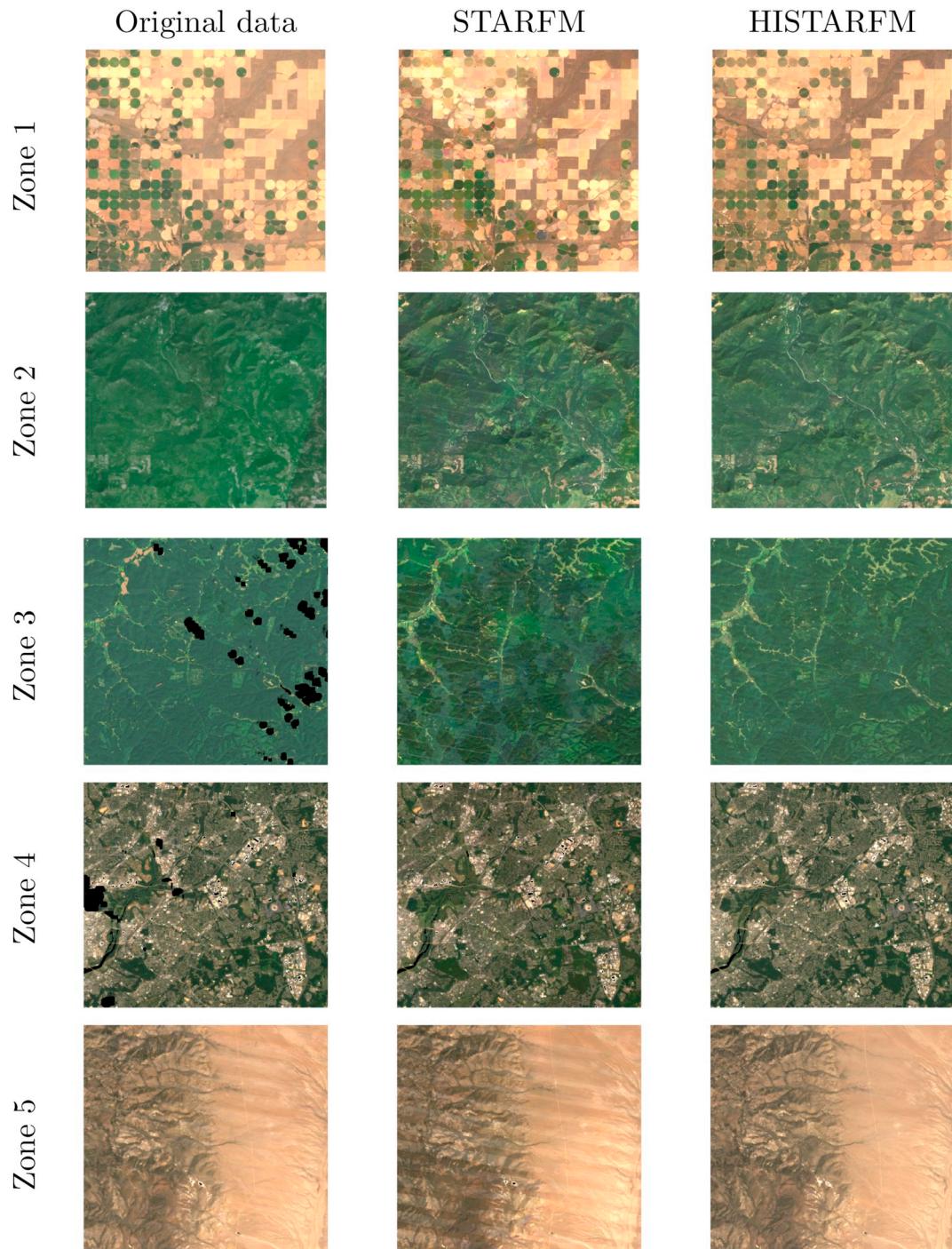


Fig. 8. Comparison of the original Landsat RGB composites and the predictions of HISTARFM and STARFM for the five selected study areas (August 2018)

made if at least one completely cloud-free and good quality observation is available for the period of interest (Hilker et al., 2009b).

Fig. 12 shows the temporal evolution of the mean residuals of the NDVI over two considered study areas (zones 4 and 5). The magnitude of the errors are in concordance with the complexity and heterogeneity of the both landscapes, being for zone 4 (very heterogeneous urban area) the double than for zone 5. In contrast with the previous analysis, where HISTARFM outperformed STARFM in all cases for the different bands, in Fig. 12 both, STARFM and HISTARFM, present low mean biases in NDVI through time. On the contrary, the standard deviation of the residuals for STARFM is always the highest, indicating higher robustness in HISTARFM predictions.

5. Discussion

5.1. Comparison with other methods and initiatives

The Web-enabled Landsat Data (WELD) project was created to provide consistent continental scale 30 m Landsat long-term data records (Roy et al., 2010). These data have been reported as urgently required for a better monitoring of the Earth system functioning and land-cover change globally and provide a high spatial resolution complementary analogue of very well established coarse spatial resolution land products from the MODIS and AVHRR sensors data streams. To our knowledge, the WELD project is the only initiative which provides

Table 4

Validation of the results for the bands B1, B2, B3, and B4 (August 2018). HI and ST refers to HISTARFM and STARFM respectively. The whole scenes have been removed and gap filled to estimate algorithms' performance with available Landsat data. The best results are highlighted in boldface.

	Band B1		Band B2		Band B3		Band B4	
	HI	ST	HI	ST	HI	ST	HI	ST
Zone 1	MBE	-0.014	-0.008	-0.012	-0.008	-0.011	-0.011	-0.03
	MAE	0.02	0.02	0.016	0.016	0.013	0.014	0.04
	RSME	0.03	0.03	0.02	0.02	0.017	0.017	0.05
	R	0.92	0.91	0.90	0.90	0.87	0.88	0.82
Zone 2	MBE	-0.011	-0.012	-0.03	-0.03	-0.016	-0.016	-0.004
	MAE	0.011	0.013	0.03	0.03	0.016	0.015	0.011
	RSME	0.012	0.014	0.03	0.03	0.016	0.017	0.014
	R	0.93	0.89	0.83	0.77	0.81	0.71	0.97
Zone 3	MBE	-0.004	-0.009	-0.007	-0.011	-0.010	-0.014	-0.018
	MAE	0.005	0.009	0.007	0.012	0.010	0.014	0.02
	RSME	0.007	0.011	0.02	0.008	0.013	0.015	0.03
	R	0.76	0.65	0.87	0.73	0.62	0.45	0.80
Zone 4	MBE	-0.001	-0.007	-0.001	-0.007	-0.002	-0.006	-0.003
	MAE	0.005	0.008	-0.005	0.009	0.007	0.010	0.012
	RSME	0.008	0.03	0.009	0.03	0.013	0.03	0.05
	R	0.96	0.65	0.96	0.69	0.96	0.77	0.94
Zone 5	MBE	0.000	-0.004	-0.002	-0.004	-0.004	-0.003	-0.001
	MAE	0.005	0.009	0.005	0.010	0.007	0.011	0.008
	RSME	0.006	0.012	0.007	0.010	0.008	0.016	0.010
	R	0.98	0.92	0.98	0.94	0.98	0.96	0.85

access to high resolution optical data at a continental scale. In this work, we pursued an identical objective but we propose a different approach which heavily relies on modern cloud computing platforms to process the massive Landsat and MODIS data sets at broad spatial scales. Moreover, instead of using per-pixel temporal compositing approach to obtain cloud free mosaics of Landsat as WELD does, we have developed and implemented in the cloud a data assimilation method which provides gap free reflectance estimations along with its uncertainties.

Previous approaches have mostly capitalized on refining and improving their data fusion algorithms to minimize the errors and noise in their final estimates. One important pitfall of these approaches is that they were not specifically designed to process efficiently large amounts of data and, very often, rely on rather complex computations that are not easily parallelizable to run in modern cloud computing platforms such as GEE. While these methods produce in general satisfactory results, they have been mostly validated with synthetic data or over rather small areas (Gevaert and García-Haro, 2015; Walker et al., 2012; Chen et al., 2015) because of the above mentioned computational cost constrain. Taking into account that high spatial resolution sensors on board of the Landsat and Sentinel platforms produce daily huge amounts of multi spectral data globally, especially designed and highly parallelizable cloud computing approaches like the one proposed in this work could be of great interest to produce continental gap filled satellite land surface reflectance data sets. To provide an example of the unprecedented computational power of GEE, the computation of gap free Landsat reflectances with HISTARFM in 2016 used 19,537 Landsat scenes (combining Landsat 7 and Landsat 8 platforms) and 92 full MODIS mosaics (combining Terra and Aqua platforms) for the study area (CONUS). Moreover, the high efficiency of the proposed method running in GEE allows to process in 2 days all Landsat optical bands for a given date covering the full CONUS, which is around processing $1.2 \cdot 10^{11}$ pixels.

The quantitative and qualitative comparison of our approach with STARFM revealed a great consistence among them (see Section 4.3). Moreover, HISTARFM outperforms STARFM for the considered bands, at the considered zones and dates. This seems to imply that most of the needed spatial operations that impede the implementation of the STARFM or similar approaches in the cloud could be circumvented using information from previous years by means of the computed Landsat climatology included in our approach. The computation of the

climatology is very simple, efficient, and it is carried out on a pixel basis but, on the other hand, involves processing huge amounts of data at a time (ten years in our case). This could be a serious limitation in the application of our method in situations when the data is not ingested in the system or the satellite data record is too short. Fortunately, the GEE platform has the whole Landsat 4, 5, 7 and 8 surface reflectance data set (starting in 1982) ready to use and it is routinely updated. Furthermore, looking into the future, this limitation will not be as important, since we will have more medium resolution data available.

Although the validation of most of the state-of-the-art approaches is very limited in space and time, we have carried out a literature review to provide some insight about the magnitude of the uncertainties and discrepancies with them and the proposed approach. This comparison with the literature has to be interpreted carefully because the accuracy and precision of the methods reported in the literature could be affected by many factors, such as the heterogeneity of land surface and the change magnitude between the target and reference data. As an example, Fu et al. (2013) modified the ESTARFM (mESTARFM) algorithm and tested its performance in three areas located in China and Canada by comparing with two other similar approaches. Correlations ranged between 0.41 and 0.89 for the band B2, 0.55–0.93 for the band B3 and, 0.67–0.97 for the band B4, while the RMSE values ranged between 0.005 and 0.02 for the band B2, 0.005–0.16 for the band B3 and, 0.005–0.03 for the band B4. Chen et al. (2015) compared four STARFM like fusion models (STARFM, SPSTFM, ISTAFM, and ESTARFM) over a wetland area in the Poyang Lake (China) and an irrigation cropland area in Coleambally (Australia). They ran the models in two different configurations. Firstly, using single MODIS/Landsat single pairs and, secondly, using two MODIS/Landsat pairs for their predictions. Their coefficients of correlation between the predicted and actual reflectance ranged from 0.30–0.95 for the band B3, 0.40–0.90 for B4, and 0.6–0.95 for B5 depending on the area and the dates analyzed. Gevaert and García-Haro (2015) proposed a new unmixing-based method which incorporated prior spectral information, the validation was carried out with both, simulated high-resolution imagery and some base images to predict in Spain. The results were also compared with the STARFM algorithm and the authors reported RMSE values of 0.012 (B1), 0.019 (B2), 0.029 (B3), 0.043 (B4), 0.057 (B5), 0.052 (B7), and correlations of 0.93 (B1), 0.92(B2), 0.94(B3), 0.82(B4), 0.87(B5), 0.91(B7). Other recent approaches proposed by Song et al. (2018) rely on a novel spatio-temporal fusion method based on deep convolutional neural networks

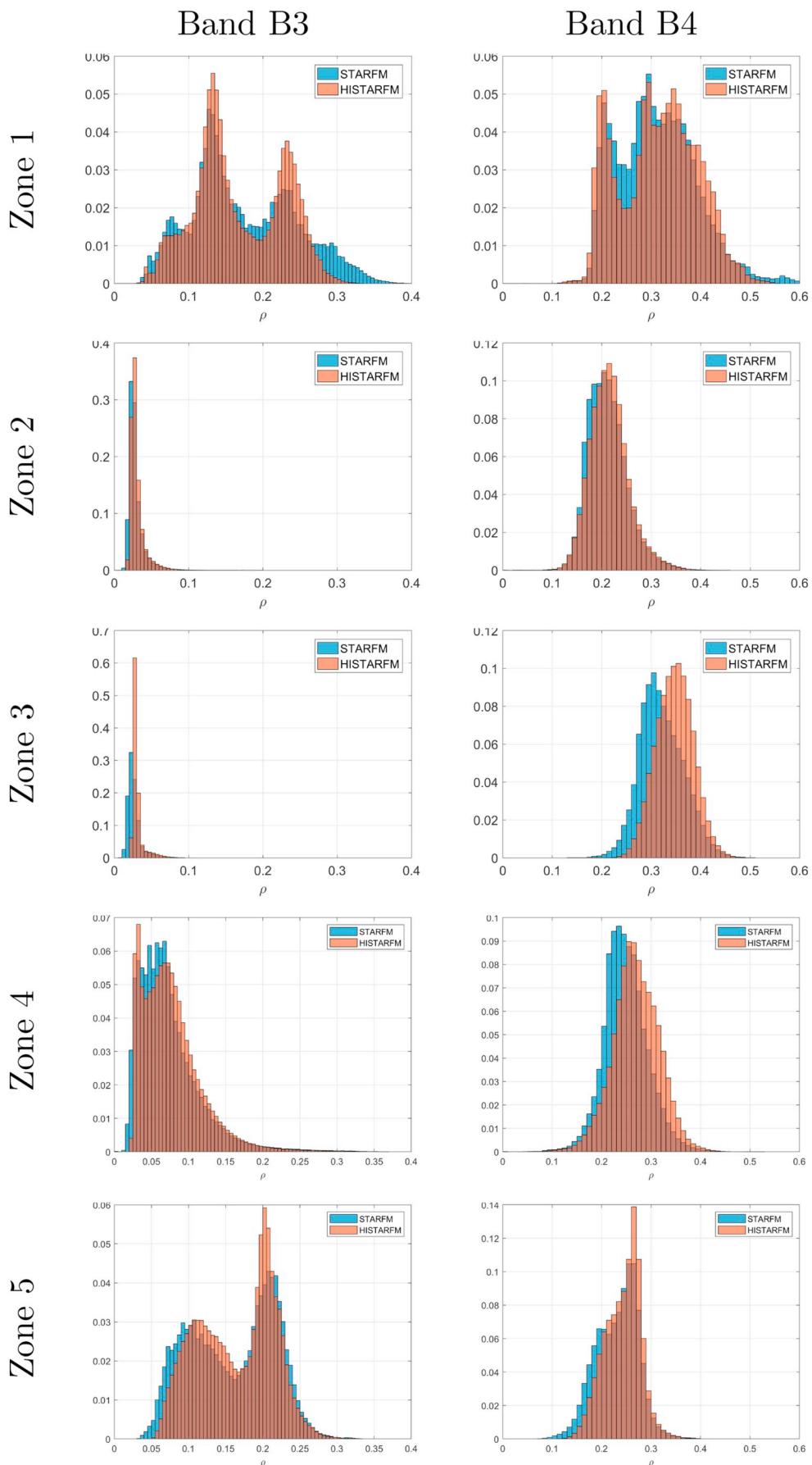


Fig. 9. Distribution of the estimated surface reflectance values over the selected sites for the red (B3) and infrared (B4) spectral bands. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

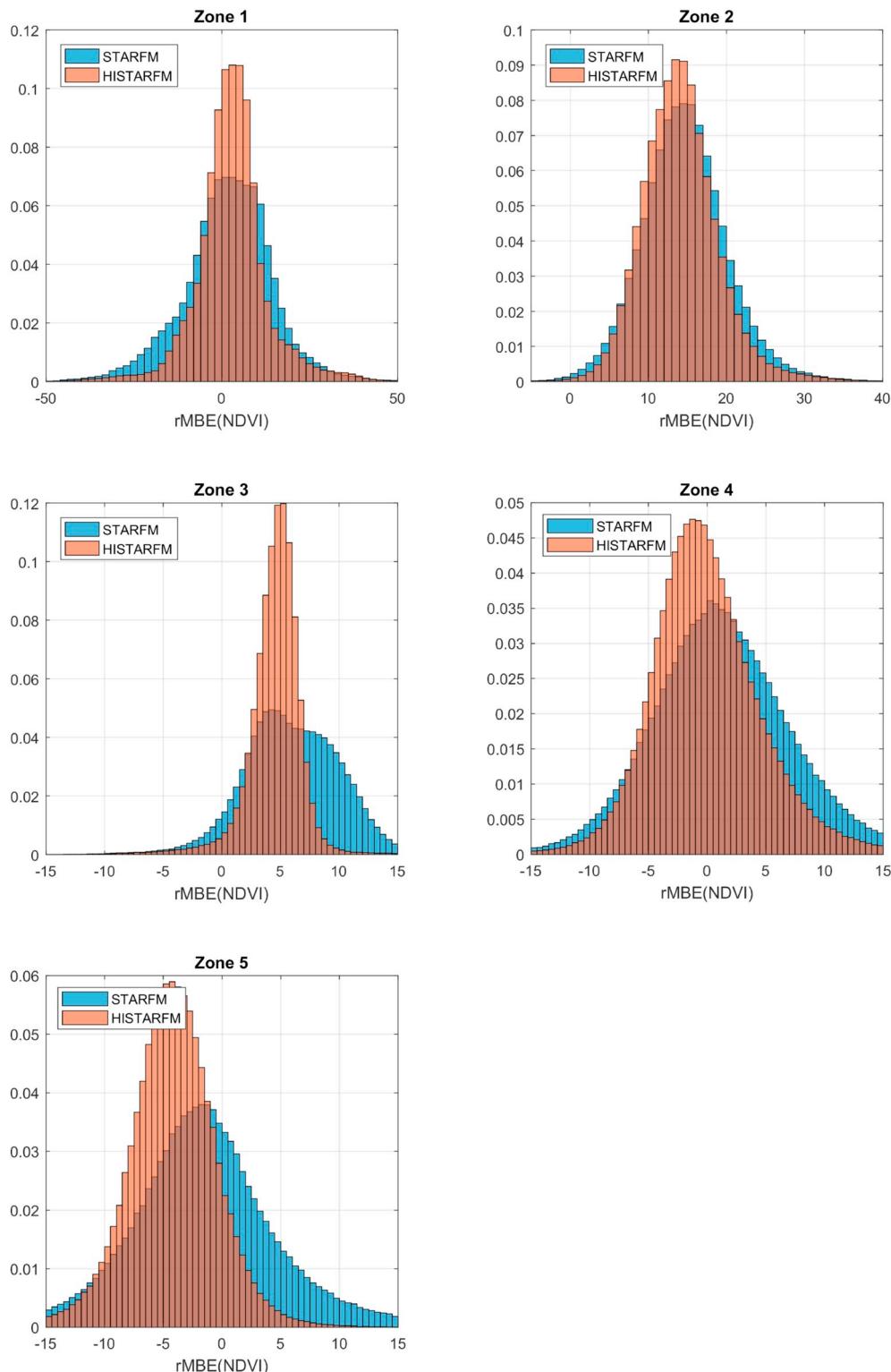


Fig. 10. Analysis of the NDVI residuals over the selected study areas by comparing the NDVI computed with actual Landsat reflectance data and gap filled bands with the STARFM and HISTARFM algorithms.

(CNNs). The evaluation of their approach in two different study areas (Coleambally and the lower Gwydir catchment in Australia) over a set of prediction dates obtained RMSE values ranging between 0.007 and 0.015 for band B1, 0.011–0.020 for band B2, 0.015–0.025 for band B3, 0.022–0.032 for band B4, 0.023–0.051 for band B5, and 0.021–0.039 for band B7. The comparison of the above mentioned statistics provided by other authors with those shown in our results section (Table 3),

suggests that the presented method compares well with state of the art gap filling and data fusion approaches despite the lack of a continental scale validation of them. This result provides a framework for a feasible operational method to provide gap free observations of high resolution multi-spectral sensors such as the Landsats or the Sentinels.

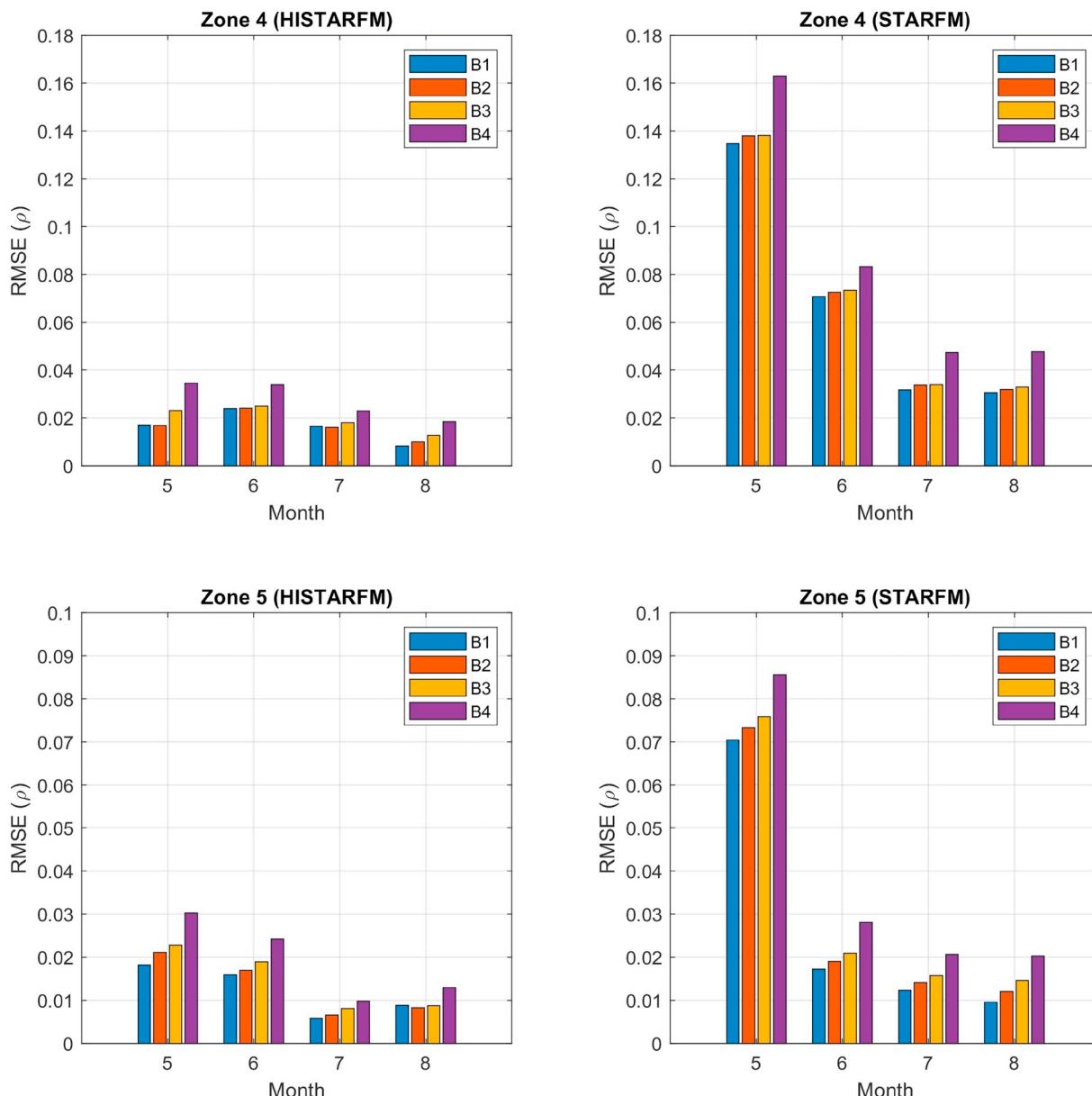


Fig. 11. Temporal analysis of the RMSE predicting Landsat reflectances with HISTARFM and STARFM algorithms over two selected study areas. Zone 4 corresponds with very heterogeneous urban area close to Washington D.C. (US), and zone 5 corresponds with a dry ecosystem area mixture of evergreen forest and shrub lands located in Nevada, (US).

5.2. Further improvements

The method presented here heavily relies on the quality of the masking and uncertainties in the measured data, in our case, the Landsat reflectance. Different authors have documented that despite the LEADPS corrections, some seasonal solar zenith variations remain in the spectral data even though the Landsat sensor has a narrow field of view (Roy et al., 2016b). For this reason, further improvements of the proposed method should include normalization methods to correct the effects of non negligible bidirectional reflectance distribution function (BRDF) undesired effects. These corrections could be implemented as a preliminary step before running the algorithm, although a more straightforward way to do it could be replacing the LEADPS data for harmonized datasets when they will be available. The Harmonized Landsat/Sentinel-2 (HLS) aims to generate a surface reflectance product by combining data from Landsat-8 and Sentinel-2 platforms while also implement a model to derive the BRDF normalization and apply it at 30 m spatial resolution (Franch et al., 2019). The FORCE (Framework

for Operational Radiometric Correction for Environmental monitoring) initiative could be also a good candidate to be used in place of the LEADPS Landsat reflectance products (Frantz, 2019). According to their authors, the FORCE is capable of generating multi-sensor analysis ready data (ARD) and is comprised of operational cloud masking and radiometric correction methods (including topographic or BRDF correction). So replacing the present Landsat data by any improved harmonized or ARD data set would be an easy and efficient way to generate better results with minimal changes in the proposed method. In the current version of HISTARFM, the algorithm has been developed and constrained to be applied band-by-band assuming the bands are independent. Further improvements of HISTARFM will take advantage of exploiting inter-band correlations for the predictions. The practical implementation of it should be further investigated, but it surely requires moving away from the actual scalar implementation. This will need more computational resources to be allocated at once but, in any case, exploiting inter-band correlations and estimating the whole spectra jointly could potentially improve the accuracy of HISTARFM

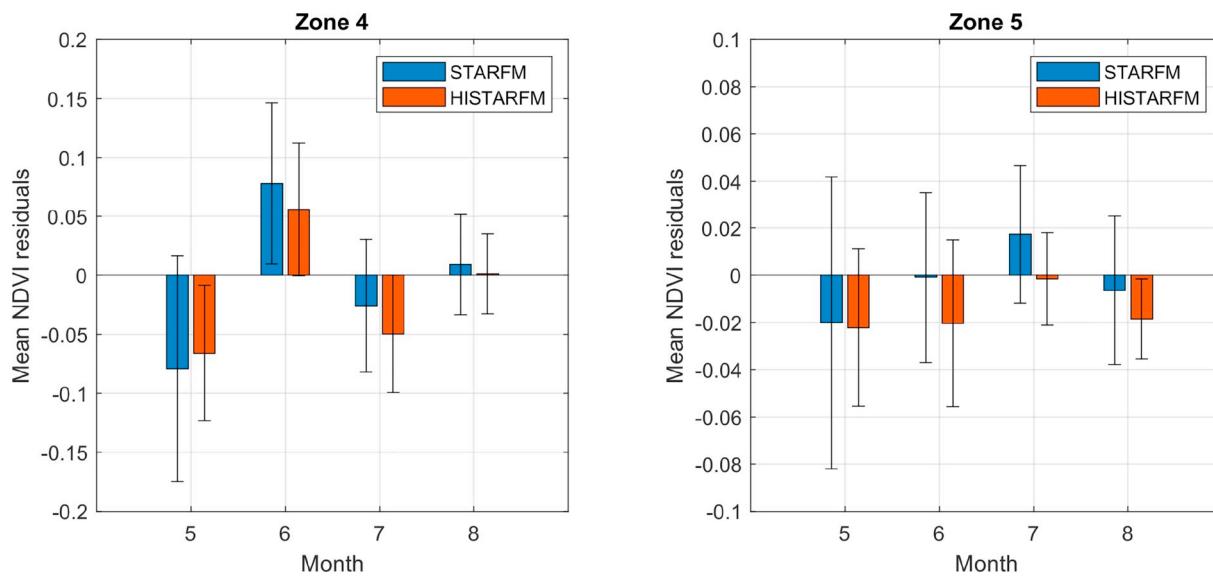


Fig. 12. Relative mean bias error for the gap filled NDVI and the observed Landsat NDVI over two study areas (zones 4 and 5).

final estimates and the calculated predicted variances.

5.3. Future applications

The Landsat program running continuously and consistently for more than forty years provides a unique opportunity to provide key higher-level products valuable for many climate studies at an unprecedented spatial resolution (Wulder et al., 2019). The provided gap filled spectral products relying on MODIS and Landsat data go a step further, and allow to create continuous time series of Essential Climate Variables (ECVs) such as the fraction of absorbed photosynthetically active radiation (FAPAR) and leaf area index (LAI), along with realistic uncertainty estimates of them. It is important to note that, because the proposed method provides reconstructed spectral reflectances, there is no need of using empirical parameterizations of vegetation indices to estimate biophysical parameters. Instead, more robust approaches which rely on the inversion of radiative transfer models could be used to retrieve the variables of interest. Additionally, because no gaps are present in the predicted time series, there is no need of compositing different years to provide annual land phenology information. This would potentially enable the creation of yearly information associated with above-ground forest biomass, burned area, crop yield prediction, and land cover change for example (He et al., 2018; Huang et al., 2019; He et al., 2019).

6. Conclusions

We have presented a new data fusion method, HISTARFM, for the smoothing and gap filling of Landsat reflectance time series at an unprecedented broad spatial scale. We implemented a bias-aware Kalman filter method in the Google Earth Engine (GEE) platform to obtain the monthly fused images at Landsat spatial-resolution. The method has been designed to take full advantage of modern cloud computing platforms, and allows to process large amounts of data by using simple and scalable operations that run optimally in the cloud. The direct validation of the proposed approach at a continental scale over 1050 sites representing common vegetation types over the study indicated the feasibility of the method. The relative mean errors (rME) remained below 1.5% for all the bands, indicating that our estimates are

unbiased. The relative mean absolute errors (rMAE) and relative root mean squared errors (rRMSE) were low to moderate, and varied significantly between bands in concordance with the LEDAPS algorithm uncertainties reported in the literature. The high degree of agreement between the validation errors and the predicted uncertainties by HISTARFM provides the foundation for using them for error propagation in future applications.

The proposed methodology is general enough to incorporate other, more advanced, nonlinear approaches beyond Kalman filtering approaches, such as particle filtering and recursive Gaussian processes. Exploitation of the generated products may have an impact in many applications for land monitoring, from phenological studies to change detection and crop yield estimation. Also, the provided gap filled reflectances go a step further and have the potential to provide continuous time series of Essential Climate Variables (ECVs) and many other key terrestrial vegetation variables at broad scales.

Acknowledgements

This research was financially supported by the NASA Earth Observing System MODIS project (grant NNX08AG87A), USDA AFRI grant 2016-67026-25067, and NASA EPSCoR, United States grant 80NSSC18M0025M. This work was also supported by the European Research Council (ERC) funding under the ERC Consolidator Grant 2014 SEDAL (Statistical Learning for Earth Observation Data Analysis, European Union) project under Grant Agreement 647423. Gap filled Landsat data processed with HISTARFM can be freely accessed by the Google Earth Engine users from https://code.earthengine.google.com/?asset=projects/KalmanGFwork/GFLandsat_V1. In addition, the data can be explored without a Google Earth Engine account from <https://almoma153.users.earthengine.app/view/explorehistarfm>. We would like to thank the Google Earth Engine developers for their support and technical advice and the anonymous reviewers for their helpful comments and suggestions.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A: Table of Symbols

Table A.5

Table with definition of symbols used in this paper in order of appearance in the text.

Symbol	Description
K_k^-	Kalman gain at timestep k
X_k^-	prior estimate of reflectance at time k
x_k	Posterior (corrected) estimate of reflectance at time step k
z_k	Landsat reflectance observation at time k
P_k^-	Error covariance of the prior estimate of reflectance at time k
P_k	Error covariance of the posterior estimate at time k
H	Observation operator relating modeled reflectances and observed reflectances
R	Landsat error covariance
β_i	Linear regression coefficients relating MODIS and Landsat reflectances for pixel i
z_i	Vector of monthly Landsat reflectances for pixel i with one element per month of selected year
$u_{i, mod}$	Vector of monthly MODIS reflectances of pixel i with one element per month of selected year.
$U_{i, MOD}$	Reflectances are resampled at 30 m resolution using a nearest neighbor algorithm.
$U_{i, k, mod}^{30}$	Augmented input matrix [$1 u_{i, mod}$] with 1 being a column vector with all elements set to 1
$u_{i, k, mod}$	Vector of MODIS reflectances at pixel i for month k spatially disaggregated at 30 m from regression operation
$U_{i, k, mod}^{30}$	Augmented input matrix [$1 u_{i, k, mod}^{30}$]
$P_{k, MOD}$	Error covariance of the spatially disaggregated MODIS reflectance
\bar{z}_k	Landsat mean reflectance of month k calculated using 10 years prior to selected year
$\bar{P}_{k, LS}$	Landsat climatological variance of month k calculated using 10 years prior to selected year
γ	Fraction of error covariance of the estimate attributable to observation bias
W_k	Error covariance of the posterior corrected unbiased reflectance estimate at time step k
T_k	Error covariance of the posterior estimate of the reflectance bias at month k
b_k^-	Prior estimate of the reflectance bias at month k
b_k	Posterior (corrected) estimate of reflectance bias at time step k
T_k^-	Error covariance of the prior estimate of the reflectance bias at month k
\bar{x}_k	Corrected and unbiased estimate of 30 m resolution reflectance of month k

References

- Amorós-López, J., Gómez-Chova, L., Alonso, L., Guanter, L., Zurita-Milla, R., Moreno, J., Camps-Valls, G., 2013. Multitemporal fusion of landsat/tm and envisat/meris for crop monitoring. *Int. J. Appl. Earth Obs. Geoinf.* 23, 132–141.
- Bradley, B.A., Jacob, R.W., Hermance, J.F., Mustard, J.F., 2007. A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite ndvi data. *Remote Sens. Environ.* 106 (2), 137–145.
- Chen, J., Jönsson, P., Tamura, M., Gu, Z., Matsushita, B., Eklundh, L., 2004. A simple method for reconstructing a high-quality ndvi time-series data set based on the svitzky-golay filter. *Remote Sens. Environ.* 91 (3–4), 332–344.
- Chen, B., Huang, B., Xu, B., 2015. Comparison of spatiotemporal fusion models: a review. *Remote Sens.* 7 (2), 1798–1835.
- Claverie, M., Vermote, E.F., Franch, B., Masek, J.G., 2015. Evaluation of the landsat-5 tm and landsat-7 etm + surface reflectance products. *Remote Sens. Environ.* 169, 390–403.
- Dee, D.P., Da Silva, A.M., 1998. Data assimilation in the presence of forecast bias. *Q. J. R. Meteorol. Soc.* 124 (545), 269–295.
- Drécourt, J.-P., Madsen, H., Rosbjerg, D., 2006. Bias aware kalman filters: comparison and improvements. *Adv. Water Resour.* 29 (5), 707–718.
- Foga, S., Scaramuzza, P.L., Guo, S., Zhu, Z., Dilley Jr., R.D., Beckmann, T., Schmidt, G.L., Dwyer, J.L., Hughes, M.J., Laue, B., 2017. Cloud detection algorithm comparison and validation for operational landsat data products. *Remote Sens. Environ.* 194, 379–390.
- Franch, B., Vermote, E., Skakun, S., Roger, J.-C., Masek, J., Ju, J., Villaescusa-Nadal, J.L., Santamaría-Artigas, A., 2019. A method for landsat and sentinel 2 (hls) brdf normalization. *Remote Sens.* 11 (6), 632.
- Frantz, D., 2019. Force+landsat+ sentinel-2 analysis ready data and beyond. *Remote Sens.* 11 (9), 1124.
- Friedland, B., 1969. Treatment of bias in recursive filtering. *IEEE Trans. Autom. Control* 14 (4), 359–367.
- Fu, D., Chen, B., Wang, J., Zhu, X., Hilker, T., 2013. An improved image fusion approach based on enhanced spatial and temporal the adaptive reflectance fusion model. *Remote Sens.* 5 (12), 6346–6360.
- Fung, C.H., Wong, M.S., Chan, P., 2019. Spatio-temporal data fusion for satellite images using hopfield neural network. *Remote Sens.* 11 (18), 2077.
- Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the blending of the landsat and modis surface reflectance: predicting daily landsat surface reflectance. *IEEE Trans. Geosci. Remote Sens.* 44 (8), 2207–2218.
- Gevaert, C.M., Garca-Haro, F.J., 2015. A comparison of starfm and an unmixing-based algorithm for landsat and modis data fusion. *Remote Sens. Environ.* 156, 34–44.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27.
- Goward, S.N., Markham, B., Dye, D.G., Dulaney, W., Yang, J., 1991. Normalized difference vegetation index measurements from the advanced very high resolution radiometer. *Remote Sens. Environ.* 35 (2–3), 257–277.
- Goward, S.N., Masek, J.G., Williams, D.L., Irons, J.R., Thompson, R., 2001. The landsat 7 mission: terrestrial research and applications for the 21st century. *Remote Sens. Environ.* 78 (1–2), 3–12.
- Gu, Y., Wylie, B., 2015. Downscaling 250-m modis growing season ndvi based on multiple landsat images and data mining approaches. *Remote Sens.* 7 (4), 3489–3506.
- Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S., Tyukavina, A., Thau, D., Stehman, S., Goetz, S., Loveland, T.R., et al., 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342 (6160), 850–853.
- He, M., Kimball, J.S., Maneta, M.P., Maxwell, B.D., Moreno, A., Beguería, S., Wu, X., 2018. Regional crop gross primary productivity and yield estimation using fused landsat-MODIS data. *Remote Sens.* 10 (3).
- He, M., Kimball, J.S., Yi, Y., Running, S.W., Guan, K., Moreno, A., Wu, X., Maneta, M., 2019. Satellite data-driven modeling of field scale evapotranspiration in croplands using the MOD16 algorithm framework. *Remote Sens. Environ.*
- Hilker, T., Wulder, M.A., Coops, N.C., Linke, J., McDermid, G., Masek, J.G., Gao, F., White, J.C., 2009a. A new data fusion model for high spatial-and temporal-resolution mapping of forest disturbance based on landsat and modis. *Remote Sens. Environ.* 113 (8), 1613–1627.
- Hilker, T., Wulder, M.A., Coops, N.C., Seitz, N., White, J.C., Gao, F., Masek, J.G., Stenhouse, G., 2009b. Generation of dense time series synthetic landsat data through data blending with modis using a spatial and temporal adaptive reflectance fusion model. *Remote Sens. Environ.* 113 (9), 1988–1999.
- Hoeting, J.A., Madigan, D., Raftery, A.E., Volinsky, C.T., 1999. Bayesian model averaging: a tutorial. *Stat. Sci.* 382–401.
- Homar, C., Dewitt, J., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N., Wickham, J., Megown, K., 2015. Completion of the 2011 national land cover database for the conterminous United States—representing a decade of land cover change information. *Photogramm. Eng. Remote Sens.* 81 (5), 345–354.
- Huang, B., Song, H., 2012. Spatiotemporal reflectance fusion via sparse representation. *IEEE Trans. Geosci. Remote Sens.* 50 (10), 3707–3716.
- Huang, J., Ma, H., Sedano, F., Lewis, P., Liang, S., Wu, Q., Su, W., Zhang, X., Zhu, D., 2019. Evaluation of regional estimates of winter wheat yield by assimilating three remotely sensed reflectance datasets into the coupled wofost–prosail model. *Eur. J. Agron.* 102, 1–13.
- Ingla, J., Vincent, A., Arias, M., Marais-Sicre, C., 2016. Improved early crop type identification by joint use of high temporal resolution sar and optical image time series. *Remote Sens.* 8 (5), 362.
- Jonsson, P., Eklundh, L., 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. *IEEE Trans. Geosci. Remote Sens.* 40 (8), 1824–1832.
- Ju, J., Roy, D.P., Vermote, E., Masek, J., Kovalsky, V., 2012. Continental-scale validation of modis-based and ledaps landsat etm + atmospheric correction methods. *Remote Sens. Environ.* 122, 175–184.
- Julien, Y., Sobrino, J.A., 2010. Comparison of cloud-reconstruction methods for time

- series of composite ndvi data. *Remote Sens. Environ.* 114 (3), 618–625.
- Kalman, R.E., 1960. A new approach to linear filtering and prediction problems. *J. Basic Eng.* 82 (1), 35–45.
- Kandasamy, S., Baret, F., Verger, A., Neveux, P., Weiss, M., 2013. A comparison of methods for smoothing and gap filling time series of remote sensing observations—application to modis lai products. *Biogeosciences* 10 (6), 4055–4071.
- Kang, S., Running, S.W., Zhao, M., Kimball, J.S., Glassy, J., 2005. Improving continuity of modis terrestrial photosynthesis products using an interpolation scheme for cloudy pixels. *Int. J. Remote Sens.* 26 (8), 1659–1676.
- Ke, Y., Im, J., Park, S., Gong, H., 2016. Downscaling of modis one kilometer evapotranspiration using landsat-8 data and machine learning approaches. *Remote Sens.* 8 (3), 215.
- Kempeneers, P., Sedano, F., Piccard, I., Eerens, H., 2016. Data assimilation of proba-v 100 and 300 m. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sensing* 9 (7), 3314–3325.
- Kobayashi, H., Dye, D.G., 2005. Atmospheric conditions for monitoring the long-term vegetation dynamics in the amazon using normalized difference vegetation index. *Remote Sens. Environ.* 97 (4), 519–525.
- Kumar, L., Mutanga, O., 2018. Google earth engine applications since inception: usage, trends, and potential. *Remote Sens.* 10 (10), 1509.
- Liang, S., 2005. Quantitative Remote Sensing of Land Surfaces. vol. 30 John Wiley & Sons.
- Liu, M., Yang, W., Zhu, X., Chen, J., Chen, X., Yang, L., Helmer, E.H., 2019. An improved flexible spatiotemporal data fusion (ifsdfa) method for producing high spatio-temporal resolution normalized difference vegetation index time series. *Remote Sens. Environ.* 227, 74–89.
- Loveland, T.R., Irons, J.R., 2016. Landsat 8: the plans, the reality, and the legacy. *Remote Sens. Environ.* 185, 1–6.
- Luo, Y., Guan, K., Peng, J., 2018. Stair: a generic and fully-automated method to fuse multiple sources of optical satellite data to generate a high-resolution, daily and cloud-/gap-free surface reflectance product. *Remote Sens. Environ.* 214, 87–99.
- Lyapustin, A., Wang, Y., Frey, R., 2008. An automatic cloud mask algorithm based on time series of modis measurements. *J. Geophys. Res.* 113 (D16).
- Lyapustin, A., Wang, Y., Korkin, S., Huang, D., 2018. Modis collection 6 maiac algorithm. *Atmos. Meas. Tech.* 11 (10).
- Malyy, M., Tekic, Z., Golkar, A., 2019. What drives technology innovation in new space?: a preliminary analysis of venture capital investments in earth observation start-ups. *IEEE Geosci. Remote Sensing Magazine* 7 (1), 59–73.
- Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Gao, F., Kutler, J., Lim, T.-K., 2006. A landsat surface reflectance dataset for north america, 1990–2000. *IEEE Geosci. Remote Sens. Lett.* 3 (1), 68–72.
- Matheron, G., 1963. Principles of geostatistics. *Econ. Geol.* 58 (8), 1246–1266.
- Moreno, Á., García-Haro, F., Martínez, B., Gilabert, M., 2014. Noise reduction and gap filling of fapar time series using an adapted local regression filter. *Remote Sens.* 6 (9), 8238–8260.
- Poggio, L., Gimona, A., Brown, I., 2012. Spatio-temporal modis evi gap filling under cloud cover: an example in scotland. *ISPRS J. Photogramm. Remote Sens.* 72, 56–72.
- Robinson, N., Allred, B., Jones, M., Moreno, A., Kimball, J., Naugle, D., Erickson, T., Richardson, A., 2017. A dynamic landsat derived normalized difference vegetation index (ndvi) product for the conterminous united states. *Remote Sens.* 9 (8), 863.
- Roy, D.P., Ju, J., Lewis, P., Schaaf, C., Gao, F., Hansen, M., Lindquist, E., 2008. Multi-temporal modis–landsat data fusion for relative radiometric normalization, gap filling, and prediction of landsat data. *Remote Sens. Environ.* 112 (6), 3112–3130.
- Roy, D.P., Ju, J., Kline, K., Scaramuzza, P.L., Kovalsky, V., Hansen, M., Loveland, T.R., Vermote, E., Zhang, C., 2010. Web-enabled landsat data (weld): Landsat etm + composited mosaics of the conterminous United States. *Remote Sens. Environ.* 114 (1), 35–49.
- Roy, D.P., Kovalsky, V., Zhang, H., Vermote, E.F., Yan, L., Kumar, S., Egorov, A., 2016a. Characterization of landsat-7 to landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote Sens. Environ.* 185, 57–70.
- Roy, D.P., Zhang, H., Ju, J., Gomez-Dans, J.L., Lewis, P.E., Schaaf, C., Sun, Q., Li, J., Huang, H., Kovalsky, V., 2016b. A general method to normalize landsat reflectance data to nadir brdf adjusted reflectance. *Remote Sens. Environ.* 176, 255–271.
- Samain, O., Roujean, J.-L., Geiger, B., 2008. Use of a kalman filter for the retrieval of surface brdf coefficients with a time-evolving model based on the ecoclimap land cover classification. *Remote Sens. Environ.* 112 (4), 1337–1346.
- Schaaf, C.B., Gao, F., Strahler, A.H., Lucht, W., Li, X., Tsang, T., Strugnell, N.C., Zhang, X., Jin, Y., Muller, J.-P., et al., 2002. First operational brdf, albedo nadir reflectance products from modis. *Remote Sens. Environ.* 83 (1–2), 135–148.
- Schmidt, G., Jenkerson, C.B., Masek, J., Vermote, E., Gao, F., 2013. Landsat ecosystem disturbance adaptive processing system (ledaps) algorithm description. Tech. rep., US Geological Survey.
- Sedano, F., Kempeneers, P., Hurt, G., 2014. A kalman filter-based method to generate continuous time series of medium-resolution ndvi images. *Remote Sens.* 6 (12), 12381–12408.
- Shen, H., Li, X., Cheng, Q., Zeng, C., Yang, G., Li, H., Zhang, L., 2015. Missing information reconstruction of remote sensing data: a technical review. *IEEE Geosci. Remote Sensing Mag.* 3 (3), 61–85.
- Shen, H., Wu, J., Cheng, Q., Aihemaiti, M., Zhang, C., Li, Z., 2019. A spatiotemporal fusion based cloud removal method for remote sensing images with land cover changes. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sensing* 12 (3), 862–874.
- Song, H., Huang, B., 2012. Spatiotemporal satellite image fusion through one-pair image learning. *IEEE Trans. Geosci. Remote Sens.* 51 (4), 1883–1896.
- Song, H., Liu, Q., Wang, G., Hang, R., Huang, B., 2018. Spatiotemporal satellite image fusion using deep convolutional neural networks. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sensing* 11 (3), 821–829.
- Thenkabail, P.S., Lyon, J.G., Huete, A., 2018. Advanced Applications in Remote Sensing of Agricultural Crops and Natural Vegetation. CRC Press.
- Vermote, E.F., Kotchenova, S., 2008. Atmospheric correction for the monitoring of land surfaces. *J. Geophys. Res.* 113 (D23).
- Vermote, E.F., Tanré, D., Deuze, J.L., Herman, M., Morcette, J.-J., 1997. Second simulation of the satellite signal in the solar spectrum, 6s: an overview. *IEEE Trans. Geosci. Remote Sens.* 35 (3), 675–686.
- Vermote, E.F., El Saleous, N.Z., Justice, C.O., 2002. Atmospheric correction of modis data in the visible to middle infrared: first results. *Remote Sens. Environ.* 83 (1–2), 97–111.
- Walker, J., De Beurs, K., Wynne, R., Gao, F., 2012. Evaluation of landsat and modis data fusion products for analysis of dryland forest phenology. *Remote Sens. Environ.* 117, 381–393.
- Wulder, M.A., Loveland, T.R., Roy, D.P., Crawford, C.J., Masek, J.G., Woodcock, C.E., Allen, R.G., Anderson, M.C., Belward, A.S., Cohen, W.B., et al., 2019. Current status of landsat program, science, and applications. *Remote Sens. Environ.* 225, 127–147.
- Zhang, C., Li, W., Travis, D., 2007. Gaps-fill of slc-off landsat etm + satellite image using a geostatistical approach. *Int. J. Remote Sens.* 28 (22), 5103–5122.
- Zhu, X., Chen, J., Gao, F., Chen, X., Masek, J.G., 2010. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sens. Environ.* 114 (11), 2610–2623.
- Zscheischler, J., Mahecha, M., et al., 2014. An extended approach for spatiotemporal gapfilling: dealing with large and systematic gaps in geoscientific datasets. *Nonlinear Process. Geophys.* 21 (1), 203–215.
- Zurita-Milla, R., Gómez-Chova, L., Guanter, L., Clevers, J.G., Camps-Valls, G., 2011. Multitemporal unmixing of medium-spatial-resolution satellite images: a case study using meris images for land-cover mapping. *IEEE Trans. Geosci. Remote Sens.* 49 (11), 4308–4317.