1 **STAIR: A generic and fully-automated method to fuse multiple sources of optical satellite**

2 **data to generate a high-resolution, daily and cloud-/gap-free surface reflectance product**

3 Yunan Luo[1], Kaiyu Guan[2,3*], and Jian Peng[1*]

4

5 1. Department of Computer Science, University of Illinois at Urbana-Champaign, IL, USA

6 2. Department of Natural Resources and Environmental Sciences, College of Agriculture, Consumer, and

7 Environmental Sciences, University of Illinois at Urbana-Champaign, IL, USA

8 3. National Center of Supercomputing Applications, University of Illinois at Urbana-Champaign, IL,

9 USA

10

11 * Corresponding authors: Kaiyu Guan (kaiyug@illinois.edu), Jian Peng (jianpeng@illinois.edu)

12

13

14 Prepared for *Remote Sensing of Environment*

15

Abstract

Surface reflectance data with high resolutions in both space and time have been desired and demanded by scientific research and societal applications. Standard satellite missions could not provide such data at both high resolutions. Fusion approaches that leverage the complementary strengths in various satellite sources (e.g. MODIS/VIIRS/GOES-R's sub-daily revisiting frequency and Landsat/Sentinel-2's high spatial resolution) provide a viable means to simultaneously achieve both high resolutions in the fusion data. In this paper, we presented a novel, generic and fully-automated method, **STAIR**, for fusing multi-spectral satellite data to generate a high-frequency, high-resolution and cloud-/gap-free data. Building on the time series of multiple sources of satellite data, STAIR first imputes the missing-value pixels (due to cloud cover or sensor mechanical issues) in satellite images using an adaptive-average correction process, which takes into account different land covers and neighborhood information of miss-value pixels through an automatic segmentation. To fuse satellite images, it employs a local interpolation model to capture the most informative spatial information provided by the high spatial resolution data (e.g., Landsat) and then performs an adjustment step to incorporate the temporal patterns provided by the high-frequency data (e.g., MODIS). The resulting fused products contain daily, high spatial resolution and cloud-/gap-free fused images. We tested our algorithm to fuse surface reflectance data of MODIS and Landsat in Champaign County at Illinois and generated daily time series for all the growing seasons (Apr 1 to Nov 1) from 2000 to 2015 at 30 meter resolution. Extensive experiments demonstrated that STAIR not only captures correct texture patterns but also predicts accurate reflectance values in the generated images, with a significant performance improvement over the classic STARFM algorithm. This method is computationally efficient and ready to be scaled up to continental scales. It is also sufficiently generic to easily include various optical satellite data for fusion. We envision this novel algorithm can provide effective means to leverage historical optical satellite data to build long-term daily, 30m surface reflectance record (e.g. from 2000 to present) at continental scales for various applications, as well as produce operational near-real time daily and high-resolution data for future earth observation applications.

42

44   1. Introduction

45   Standard satellite missions always have to live with the trade-offs between spatial resolution and temporal

46   frequency. High spatial resolution and high revisiting frequency cannot be achieved simultaneously at one

47   satellite sensor. For example, coarse-resolution satellite data have daily (e.g. MODIS and VIIRS) or even

48   sub-hourly sampling frequency (e.g. GOES-R), but their spatial resolution is low (500m-a few km); high-

49   /medium-resolution satellite data have much finer spatial granularity (e.g. Landsat: 30m, Sentinel-2:

50   10m), but the sampling frequency is low (~multiple days to weeks). On the other hand, societal

51   applications have a pressing need for more satellite data that have high resolutions in both space and time

52   (Hossain, 2015; Henson, 2008; Denis et al., 2017). Specifically for agricultural related applications, field

53   level information (a few meters to tens of meters) is needed to guide farmers' precision management

54   (Mulla 2013), while rapid and subtle changes of crop growth condition or crop disturbance after extreme

55   weather events require high frequency images (e.g. daily) to capture (Gao et al., 2017). Similar

56   requirements are called for ecosystem monitoring and conservation (e.g. disturbance and change

57   detection) (Hansen et al., 2013), urban applications (Stone et al., 2001; Imhoff et al., 2010), and economic

58   surveillance (Jean et al, 2016), etc.

59

60   The above dilemma can be alleviated with the increased availability of high-/medium-resolution satellite

61   data (e.g. Sentinel-2 and commercial satellites), but data gaps inevitably still exist and remain challenging

62   to be filled. For example, though an increased sampling frequency can help to a certain extent, pervasive

63   tropical cloudiness will continue to affect image quality, and the overlaps of rainy season and crop-

64   growing season in rainfed agriculture (which produce ~75% of total global staple crop) (Portmann et al.,

65   2010) will continue to pose challenges for optical sensors to get frequent and clear-day images for crop

66   monitoring. Intuitively, integrating all the available and complementary information from multiple

67   sources of satellite data is a natural way to maximize information contents and reduce data gaps. Thus

68   methods to integrate various satellite data for a fused product that can have both high resolutions in space

69   and time have been developed.

70

71   Previous satellite fusion methods have achieved a good success in their applications. The most widely

72   known and used algorithm, the Spatial and Temporal Adaptive Reflectance Fusion Model (STARFM)

73   developed by Gao et al. (Gao et al., 2006) and its variates (e.g. ESTARFM) (Zhu et al., 2010) have

74   effectively generated Landsat-MODIS fusion products for local or regional applications in crop

75   monitoring (Gao et al., 2017), ecosystem disturbance detection (Hilker et al., 2009; Gaulton et al., 2011),

76   field-level disaggregation of evapotranspiration (Cammalleri et al., 2014). However, there exists large

77   space for improvements for more advanced and generic fusion approaches. STARFM and its successor

78   ESTARFM share the same fusion paradigm (Gao et al., 2015). Taking as input the MODIS image of the

79   target date and one or two clear-day Landsat-MODIS image pairs of other matching dates (preferably

80   dozens of days apart from the target date), STARFM (or ESTARFM) predicts a fine-resolution image of

81   the target date using a weighted neighborhood voting process. This paradigm mainly relies upon the

82   availability of clear Landsat and MODIS images on matching dates, which, however, is not often the case

83   in practice. Satellite images, especially Landsat, often contain missing values due to cloud contamination

84   or un-scanned pixels (e.g. failure of Scan Line Corrector (SLC-off) in Landsat 7), which makes it difficult

85   to search Landsat-MODIS image pairs on match dates that are temporarily close to the target date. This

86   fact greatly limits the applicability of methods that share the same spirit as STARFM. For example, when

87   a single pixel was cloud contaminated in both of the two input Landsat images, STARFM would fail to

88   predict the value of this pixel in the fine-resolution output image. The only feasible solution available

89   currently to alleviate this issue is to continue searching the closest possible Landsat-MODIS image pair of

90   a matching date, even if it is months away from the target date. This solution seeks the data availability at

91   the cost of prediction accuracy, as image pairs that are months apart from the target date would provide

92   less accurate information for the data fusion process.

93

94   Another challenge of Landsat-MODIS fusion is the need of missing-value pixels imputations. Due to

95   failure of the scan-line corrector (SLC) of the Landsat 7 Enhanced Thematic Mapper Plus (ETM+) sensor

96   in 2003, roughly 22% pixels remained to be unscanned as strips with missing values in Landsat 7 images,

97   which are referred to as SLC-off images (Arvidson et al., 2006; Ju & Roy, 2008). The un-scanned pixels

98   have become one of the major limitations of the applicability of Landsat 7 images thereafter due to the

99   missing information in areas corresponding to the un-scanned pixels. To alleviate the impact imposed by

100   this spatial discontinuity on the utility of Landsat 7 data, a number of methods have been developed to fill

101   the un-scanned gaps in the SLC-off images (Pringle et al., 2009; Chen et al., 2011; Zhu et al., 2012; Zeng

102   et al., 2013; Yin et al., 2017). These methods fill the gaps in SLC-off images by using one or two clear

103   images (reference images) of the same location acquired at different dates. The major limitation of these

104   methods is that it often requires much user effort, since these methods can only fill a single SLC-off

105   image at a time, and require users to manually select and specify one or two clear reference images for the

106   methods, which makes it difficult to couple these methods with the Landsat-MODIS fusion methods (e.g.,

107   STARFM) to generate gap-/cloud-free products in an effort-light way.

108

109   To address the above challenges, in this study we propose STAIR (satellite data integration), a novel and

110   generic fusion method, and test it for the Landsat-MODIS fusion. The major advance of STAIR is its data

111   integration ability. STAIR can take a time series that consists any number of Landsat-MODIS image

112   pairs, systematically integrate the information available in the time series for missing-pixel imputation,

113   and automatically determine the weight that each pair should contribute to the data fusion for the target

114   date. By integrating the data in the time series, in principle our method is able to fuse a fine-resolution

115   image without any unfilled pixels. Although typically considered as medium-resolution data, Landsat

116 (30m resolution) is regarded as the high-resolution data in this work, as it has a finer spatial resolution

117 than MODIS (500m resolution).

118

119 In this paper, we will first describe the methodology of STAIR and then demonstrate the applicability of

120 STAIR on gap-filling and fusing the surface reflectance data of Landsat and MODIS for the Champaign

121 County, Illinois, from 2000 to 2015, to generate a 30 meter, daily and cloud-/gap-free product for the six

122 spectral bands: red, green, blue, near-infrared (nir), shortwave infrared 1 (swir1), and shortwave infrared-

123 2 (swir2).

124 2. Materials and Methods

125 2.1 Methodology Overview

126 We present **STAIR**, a generic method to fuse multiple satellite imageries (surface reflectance) into daily

127 high-resolution products. STAIR takes time series image data from two or more sources of satellite data

128 as input, imputes pixels with missing value in images (Section 2.2, Figure 1), and generates high-

129 resolution and high-frequency products through a fusion process (Section 2.3, Figure 2). By integrating

130 and fusing multiple different types of satellite data, our method simultaneously captures both the

131 temporary changing trends and spatial texture patterns provided by each type of data.

132 2.2 Missing-pixel imputation

133 The utility of Landsat 7 SLC-off data was largely impacted by the spatially discontinuity imposed by the

134 un-scanned gaps. In addition, Landsat/MODIS images of frequently cloudy areas often have a large

135 fraction of missing-value pixels masked out by the cloud mask, and contain only a small portion of pixels

136 with unmasked values. The gap filling of the un-scanned and cloudy pixels thus not only help recover the

137 complete view of a Landsat/MODIS image, but also facilitate downstream applications and analysis, such

138   as data fusion of Landsat-MODIS images. In this subsection, we will describe our approach for imputing

139   the missing-value pixels due to un-scanned gaps or clouds in the Landsat/MODIS images.

140

141   **2.2.1 Step 1: Temporal interpolation through a linear regression**

142   One naive approach for filling the gaps in Landsat images is to use a linear regression method to

143   interpolate the missing values in a Landsat image based on the available values of the same areas in other

144   temporarily close Landsat images. The rationale behind this approach is that, in a reasonable short time

145   frame, the values of surface reflectance tend to change linearly. This property of local linearity enables

146   one to derive the missing surface reflectance values in a SLC-off image by linearly interpolating the

147   values of another one or two Landsat images on other dates (referred to as reference images), where the

148   surface reflectance of the same region are not missing. If there are more than one reference images

149   available temporarily close to the SLC-off image within a short time frame, one can fit a linear regression

150   model to obtain a more robust estimation of the missing value in the SLC-off image. Since strips in

151   Landsat 7 SLC-off images are generally offset in the time series over the same region, and Landsat 5 and

152   8 are functioning without un-scanned gaps, it is likely to find available gap-free pixels in other dates that

153   are temporarily close to the date where the image is SLC-off. This makes the linear interpolation a

154   feasible approach for gap filling. Specifically, let $L(p_g, t_0)$ be the surface reflectance value at pixel p_g in

155   the un-scanned strips of a SLC-off image at date t_0, and $L(p_g, t_1), \ldots, L(p_g, t_m)$ be the pixels at the same

156   position in other m gap-free reference images at dates $t_1, \ldots, t_m$, respectively. To derive the value

157     $L(p_g, t_0)$, we use a linear regression model, where the missing value $L(p_g, t_0)$ is filled as

$$L_{linear}(p_g, t_0) = at_0 + b$$

158   where a and b are the slope and intercept of the linear regression model, respectively, fitted by values

159     $L(p_g, t_1), \ldots, L(p_g, t_m)$. For the simplicity of notations, here we assume the date t_j ($j = 0,1,\ldots,m$) is

160   represented by the day of the year so that it can be used for numeric computation.

161

162 Despite its simplicity and effectiveness, the linear interpolation based approach does not always give the

163 accurate gap filling of the missing pixels, possibly due to that i) the time series of surface reflectance may

164 change rapidly in some time frame throughout the year and loses the local linearity, and ii) the systematic

165 bias of the reference images may lead to inaccurate linear interpolation, causing the derived value of

166 missing values in the SLC-off images to be over- or under-estimated. To produce high-quality gap-filled

167 images as input for our fusion algorithm, however, the magnitude of filled pixel values should match that

168 of the original pixel values out of gaps, without displaying visually noticeable stripes. A key observation

169 is that although inharmonic with the original pixels out of the gaps, the filled gaps have captured the

170 correct texture patterns. Therefore, we propose to correct the biased color scale of the filled pixels using a

171 neighborhood correction method.

172

173 **2.2.2 Step 2: Neighborhood correction: Global-average correction Vs Adaptive-average correction**

174 The rationale behind our neighborhood correction method for gap filling is that the relative relationship

175 between a pixel and a neighborhood pixel of it remains roughly the same across different dates in the time

176 series of Landsat data, especially in a short time frame. For example, let $L(p_g, t_0)$ be the value at a pixel

177 p_g in the un-scanned strips of a SLC-off image at date t_0, and $L(p_f, t_0)$ be the value at a pixel p_f in the

178 same SLC-off image, which is a neighborhood pixel of p_g and out of the strips. Similarly, let $L(p_g, t_1)$

179 and $L(p_f, t_1)$ be the values of pixels on the same positions in a gap-free image at another date t_1. Note

180 that here $L(p_g, t_1)$ is not a missing value. Our assumption is that if the two dates t_0 and t_1 are within a

181 short time frame (e.g., less than 2-3 weeks), the relative relationship between pixels p_g and p_f will remain

182 roughly the same. That is,

$$L(p_g, t_0) - L(p_f, t_0) \approx L(p_g, t_1) - L(p_f, t_1)$$

183 Based on this assumption, we can first compute the relative difference between the pixel p_g and p_f in the

184 clear gap-free image, and use this relative difference to correct the value $L_{linear}(p_g, t_0)$ derived by the

185 linear regression model. To make the correction more robust and reliable, we do not compute the relative

186 difference using values at two single pixels. Instead, we calculate the relative difference between two

187 "communities" of pixels:

$$\delta = \underline{L}(C_g, t_1) - \underline{L}(C_f, t_1)$$

188 where C_g is the set of all pixels that lie in the gaps of the SLC-off image, C_f is the set of all pixels that are

189 located out of the gaps; $\underline{L}(C_g, t_1)$ is the average value of pixels in the set C_g and $\underline{L}(C_f, t_1)$ has the similar

190 meaning. The global average makes the estimation of the relative difference more robust to the inherent

191 noise or error of the data. The global-average correction of $L_{linear}(p_g, t_0)$ is subsequently computed as

$$L_{global}(p_g, t_0) = L_{linear}(p_g, t_0) + \underline{L}(C_f, t_0) + \delta$$

192 Where $\underline{L}(C_f, t_0)$ is the average value of pixels in the set C_f of the SLC-off image at date t_0. We refer to

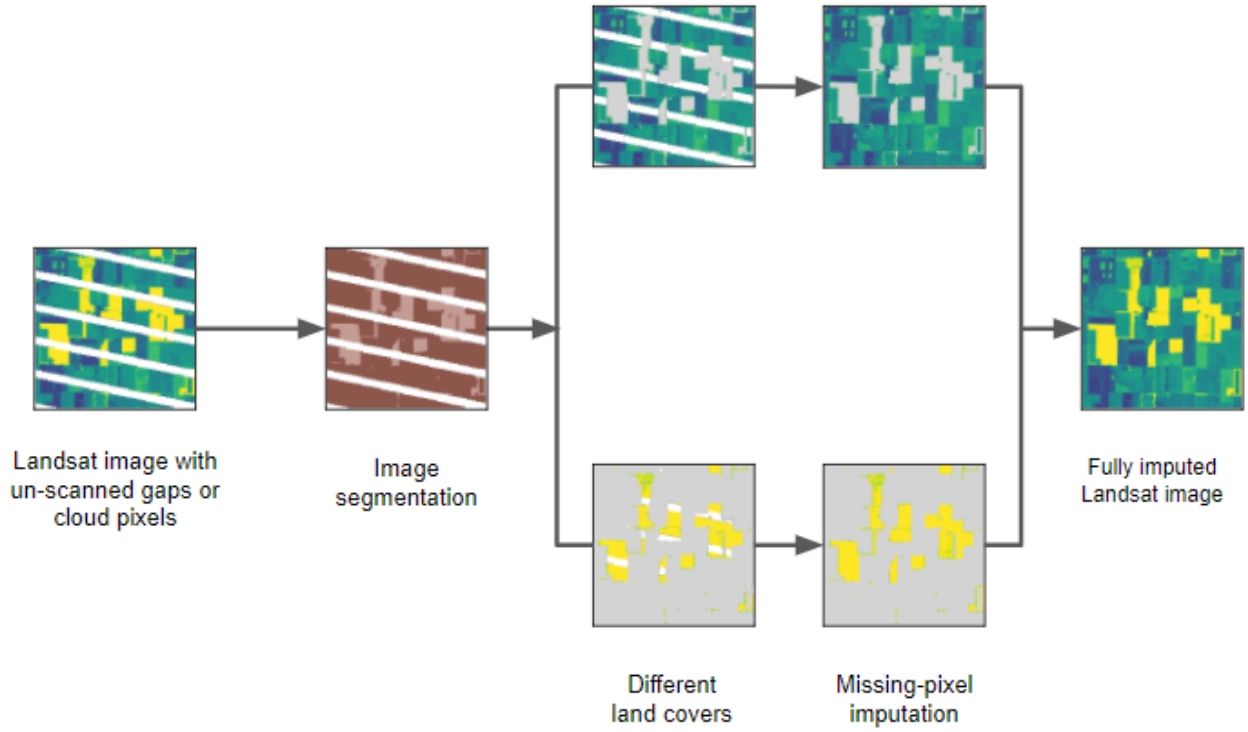193 this approach as imputation with global-average correction.

194

195

196 **Figure 1. Missing-value pixels imputation process of STAIR.** To impute the missing-value pixels, STAIR first partitions the

197 image into multiple segments, in which each segment contains one type of homogeneous pixels. Missing-value pixels are

198 imputed separately in each segment, using an adaptive-average correction process.

199

200   Different types of land cover may exhibit different trends of change in the time series of surface

201   reflectance. A global correction on the derived values for heterogeneous pixels may not lead to corrected

202   values that simultaneously compatible with all types of land cover. For instance, we observed that at the

203   end of the growing season, the global correction on the derived values accurately corrected the values for

204   one crop type, but the corrected values for another crop type were still visually incompatible, in part due

205   to different and rapid changing patterns of the two crop types in the time series. To take into account the

206   different changing patterns of heterogeneous pixels, we further extend our gap-filling algorithm to

207   adaptively correct the filled values for different land cover types separately. For this purpose, we first

208   partition the SLC-off image into multiple segments, where each segment contains the pixels

209   corresponding to one type of homogeneous pixels (Figure 1). Our gap-filling algorithm can automatically

210   choose the number of segments that best describes the number of types of homogeneous pixels in an

211   image. Specifically, we run the k-means clustering algorithm (Lloyd, 1982) to partition the image into

212   multiple segments, such that pixels of the same homogeneous type are grouped together. To automatically

213   determine the number of segments (N_s) that best explains the land cover heterogeneity in an image, we

214   used an index called gap statistic derived in (Tibshirani et al., 2001), which quantifies the dispersion of

215   the segmentation results of the image with respect to a null reference distribution (i.e., a random image

216   with no obvious clustering). A higher gap statistic generally indicates a clustering result that better

217   captures the group patterns in the data. In this work, we run the k-means clustering algorithm and

218   compute the gap statistic for the number of segments N_s varying from 2 to 8 (Figure S5). The

219   segmentation result with the highest gap statistic value is selected as the optimal segmentation of the

220   image. We found the optimal value of N_s is mostly 2, 3 or 4 for the study area in this work. After the

221   image segmentation, we similarly compute the relative difference for each segment c as follows:

$$\delta_c = \underline{L}(C_{g,c}, t_1) - \underline{L}(C_{f,c}, t_1)$$

222   where $C_{g,c}$ represents the set of pixels in gaps that are classified to segment c, and similar meaning for

223    $C_{f,c}$. Next, we adaptively correct the each pixel $p_{g,c}$ that in gaps of every segment c,

$$L_{adap}(p_{g,c}, t_0) = L_{linear}(p_{g,c}, t_0) + \underline{L}(C_{f,c}, t_0) + \delta_c$$

10

224   We refer to this approach as imputation with adaptive-average correction.

225 2.3 Fusion

226   We describe our fusion framework for integrating Landsat and MODIS data into a synthetic product with

227   both high spatial and high temporal resolutions (Figure 2). Taking the time series of Landsat and MODIS

228   images observed on matching days as input, our algorithm captures the fine-grained information provided

229   by Landsat images, and utilizes the daily MODIS images to further correct the temporal changing patterns

230   in the predicted time series of images. Previous fusion algorithms (e.g., STARFM (Gao et al., 2006)) can

231   take only one or two Landsat-MODIS pairs as input at a time, or process multiple pairs in a time-

232   consuming, pair-by-pair way with user's manual selection, in which the information used to fuse the

233   target image is often limited to the one or two input pairs. In contrast, our algorithm can process a time

234   series that consists of a much larger number of Landsat-MODIS pairs (e.g., a time series of images that

235   covers the growing season in a year), fully exploit the spatial and temporal information in the time series,

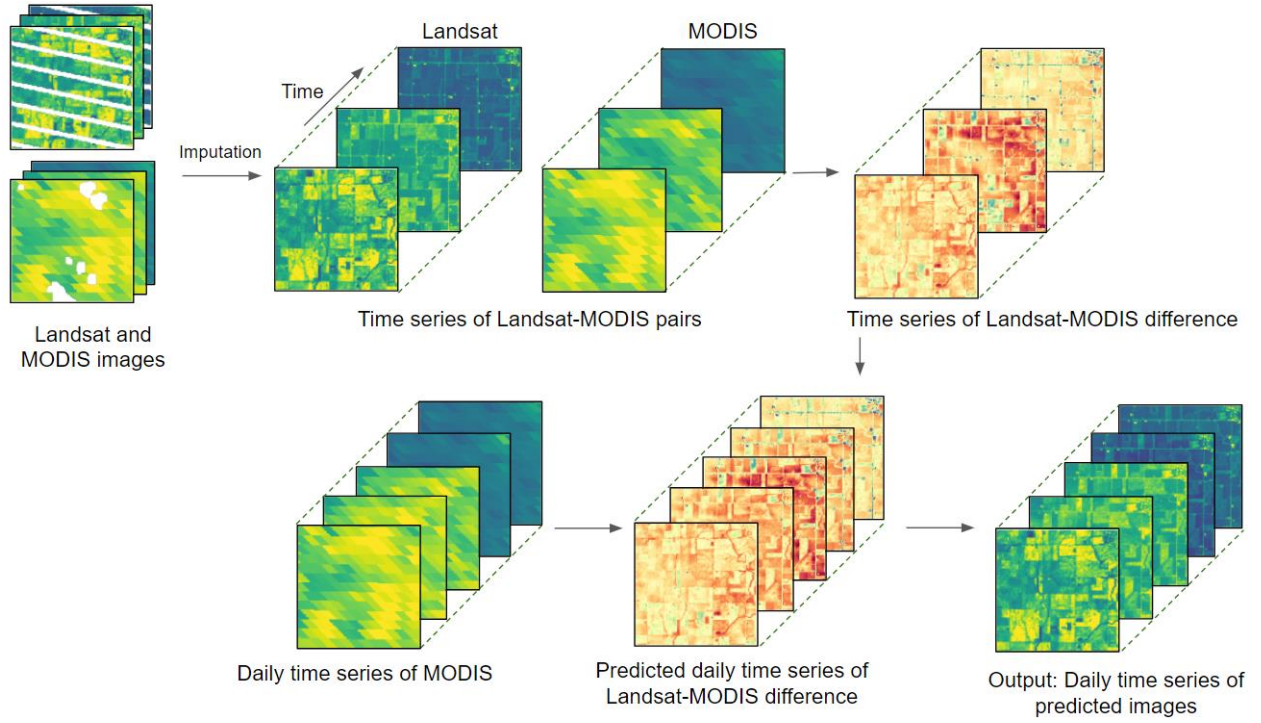236   and minimize the user's effort.

237

238

Figure 2. Schematic overview of the fusion method of STAIR. Taking Landsat and MODIS images as input, STAIR first

imputes missing-value pixels due to un-scanned gaps or clouds in the images, and then constructs a time series of Landsat-

MODIS pairs. STAIR models and predicts the Landsat-MODIS difference at a daily level, capturing the fine-resolution spatial

information. The input daily time series of MODIS data is then used to further correct the temporal patterns in the predicted time

series of Landsat-MODIS difference, generating daily time series of predicted images.

The goal of our fusion framework is to fuse Landsat and MODIS data into images with both high-

resolution spatial information and frequent temporal coverage for n dates t_i $(i = 1,2,..,n)$, and there are

k matching pairs of Landsat and MODIS images that acquired on the same dates T_j $(j = 1,2,..,k)$ among

the n dates (note that previous methods such as STARFM only supports $k = 1$ or 2). Here, we assume the

MODIS image M has been aligned and super-sampled to the resolution and bounds of the Landsat image

L, and hence has the same image size, pixel resolution and projected coordinate system. Each Landsat-

MODIS matching pairs was co-registered by first shifting the Landsat image by up to 20 pixels (since one

MODIX pixel approximately corresponds to one 17*17 Landsat pixel) in each of the four cardinal and

four ordinal directions. The optimal offset that maximizes the correlation between the Landsat and

254 MODIS images was then used to co-register the Landsat-MODIS pair. The MODIS data $M(x, y, t_i)$ is

255 also available for each date t_i $(i = 1,2,..,n)$ where we will predict the fine-resolution image.

256

257 Specifically, consider one Landsat image L and one MODIS image M that are acquired on the same date

258 T_j, the relationship for a homogeneous pixel between the Landsat image and the MODIS image can be

259 modeled as

260 $$L(x, y, T_j) = M(x, y, T_j) + \epsilon(x, y, T_j),$$

261 where (x, y) is the location of the aligned pixel of Landsat and MODIS data, T_j is the acquisition date for

262 Landsat and MODIS data, and $\epsilon(x, y, T_j)$ is the difference between Landsat and MODIS data of position

263 (x, y), due to measurement errors, solar/viewing angle geometry, noise and/or systematic biases (Zhu et

264 al., 2010). Suppose we aim to predict a fine-resolution image on date t_p where only the MODIS data

265 $M(x, y, t_p)$ is available but the Landsat data is unknown. The predicted image $I(x, y, t_p)$ can be

266 represented as $I(x, y, t_p) = M(x, y, t_p) + \Delta(x, y, t_p)$, where $\Delta(x, y, t_p)$ is the difference between the

267 MODIS image M and the predicted fine-resolution image I of position (x, y) at date t_p. To obtain the

268 difference $\Delta(x, y, t_p)$ of each position (x, y) at date t_i that we want to predict the fine-resolution image,

269 we integrate all the k available matching pairs of Landsat and MODIS data on dates from T_1 to T_k, and

270 calculate the accurate difference between Landsat and MODIS data for each date T_j $(j = 1,..,k)$ as

271 follows

272 $$\epsilon(x, y, T_j) = L(x, y, T_j) - M(x, y, T_j).$$

273 We then linearly interpolate the difference $\epsilon(x, y, T_j)$ into each date t_i that we want to predict a fine-

274 resolution image, and hence obtain an estimated difference $\Delta(x, y, t_i)$ for all positions (x, y) each date t_i.

275 Intuitively, the estimates $\Delta(x, y, t_i)$ of image differences capture the most informative and high-resolution

276 spatial information provided by Landsat data. To obtain the prediction of the fine-resolution image

277 $I(x, y, t_i)$, we use MODIS data to further correct the spatial information encoded in $\Delta(x, y, t_i)$ as follows

$$I(x, y, t_i) = M(x, y, t_i) + \Delta(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{t_i})$$

278 This step can be thought as an adjustment step that incorporates the high-frequency temporal patterns

279 provided by MODIS data. The predicted images $I(x, y, t_i)$ $(i = 1, 2, .., n)$ thus capture both high-

280 resolution spatial and temporal information. After the fusion process, a data quality flag layer is also

281 produced to indicate the confidence level of fusion results for each pixel in the fused image (Table S3 and

282 Table S4).

283 ## 2.4 Dataset

284 We build a dataset covers Landsat and MODIS images of Champaign County, IL, in multiple years to

285 evaluate the fusion performance. To perform a comprehensive evaluation to assess the ability of our

286 method under different conditions (e.g., time points and satellite sensors), the Landsat and MODIS

287 images in the dataset were selected to cover a wide temporal range, i.e. growing seasons (Apr 1-Nov 1) of

288 years from 2000 to 2015, and cover six spectral bands (i.e. red, green, blue, nir, swir1, and swir2).

289 Specifically, we collected surface reflectance data of all the Landsat data covering this period (i.e.

290 Landsat-5 TM, Landsat-7 ETM+, Landsat-8 OLI) (Masek et al., 2006; Roy et al., 2014), and MODIS

291 MCD43A4 Nadir BRDF-Adjusted Surface Reflectance (NBAR) products (Schaaf et al., 2002; Wang et

292 al., 2015) for the year 2000 to 2015 that cover Champaign County, IL. Each of the three Landsat products

293 has 30 m spatial resolution but low temporal resolution and (16-day revisit cycle), and are often cloud-

294 contaminated. While the MODIS product has a more frequent revisit cycle (daily) but a coarse spatial

295 resolution is with (500 m) resolution. Both Landsat and MODIS surface reflectance data have gone

296 through radiance calibration and atmospheric corrections and only clear-day pixels are used, i.e. the

297 fmask algorithm (Zhu & Woodcock, 2012) is applied to Landsat to remove cloudy pixels, and only "good

298 quality" flagged pixels in MODIS data are used for fusion. MODIS MCD43A4 data have also been

299 adjusted and normalized to be nadir view at local solar noon time, using the BRDF correction (Strahler &

300 Muller, 1999). The standard surface reflectance data of Landsat has not been BRDF-corrected, but due to

301 the small field of view in Landsat ($\pm7.5°$ from nadir), Landsat can be largely regarded as in nadir view

302   (Gao et al., 2006). Landsat and MODIS data are co-registered and resampled before implementing our

303   method.

304

305   To evaluate the performance of fusion results, we selected Landsat images that were captured during

306   growing seasons (Apr 1 to Nov 1) in years from 2000 to 2015 and with the percentage of non-cloudy

307   pixels greater than 60%. We then sampled multiple image patches with size 200×200 pixels uniformly at

308   random from each of these images, such that every pixel was covered once on average. These image

309   patches were held out as ground truth (referred as ground truth images) for evaluating fusion

310   performance.

311

312   We compared our method with STARFM (Gao et al., 2006), the most widely used fusion algorithm for

313   Landsat-MODIS images. We used the "two-pair" option of STARFM, in which two Landsat-MODIS

314   pairs (referred to as reference image pairs) and the MODIS image at the predicted date were used to

315   produce a fine-resolution image at the predicted date. STARFM does not provide automatic routine to

316   select the two reference pairs for fusion, we thus selected two Landsat-MODIS pairs that were within 20

317   days apart from the ground truth image and with the percentage of non-cloudy pixels greater than 60%.

318   Ground truth images that were not accompanied by two reference image pairs satisfying this criterion

319   were filtered out in our experiments of evaluation. The temporal distance (20 days) between the reference

320   image pairs and the ground truth image we set for STARFM was much smaller than the one has been

321   used in previous works of STARFM (Gao et al., 2006).

3. Results

3.1 Imputation of un-scanned and cloudy pixels

The imputation of missing pixels (e.g., cloudy pixels or un-scanned gaps in SLC-off images) can facilitate the downstream Landsat-MODIS fusion process. In this section, we show the results of missing-pixel imputation in our fusion pipeline.

The most straightforward way to fill the missing pixels in a Landsat image is to employ a linear interpolation approach based on the available pixel values of the same area in Landsat images of other temporarily close dates. The underlying assumption is that the surface reflectance changes linearly in a short time frame. However, this may not hold for the period in which the surface reflectance changes rapidly and thus exhibits a nonlinear trend. For example, we linearly interpolated the missing pixels in the red band of a SLC-off image on August 28th, 2010 (day of year 240) which was at the end of the growing season (Figure 3a). The surface reflectance of red band typically drops rapidly in this month in a nonlinear way. As expected, we found that although the linear interpolation completely filled all the missing values in the un-scanned gaps, the derived pixels are apparently incompatible with other pixels out of the gaps, displayed at visually different color scales (Figure 3b).

We then tested the imputation with global-average correction approach, in which we first filled the missing values with a linear regression model, and then corrected the bias in filled values based on the global average of available values in the SLC-off image. We observed that the imputation with global-average corrected the filled values in one type of land cover accurately, but noticeable gaps still remained in the other type of land cover (Figure 3c). The reason is that the two types of land cover have different changing patterns in the time series of surface reflectance. Therefore, the global-average correction

345   approach, which was unaware of the subtle differences in the changing patterns of the two types of land

346   cover, may lead to accurately corrected values for both of the two land covers simultaneously.

347

348   Next, we applied our proposed imputation with adaptive-average correction approach. We found that our

349   imputation algorithm with adaptive-average correction successfully addressed this issue (Figures 3c and

350   3d, red boxes). By detecting and decoupling the two types of land cover and correcting filled values

351   separately, the algorithm produced a unified image, in which all the filled pixels were accurately

352   corrected to the compatible color scale and the previously visually-noticeable gaps were eliminated. This

353   demonstrated the flexibility of our method on imputing missing values for heterogeneous landscapes.
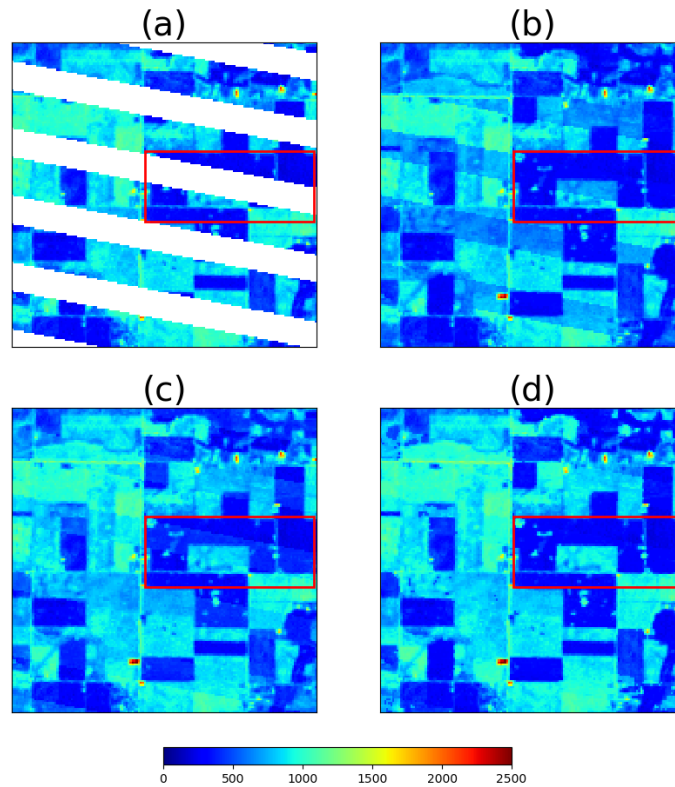
354

355

356

357   **Figure 3. Examples of missing-value pixels imputation using adaptive-average correction.** (**a**) Original SLC-off Landsat 7

358   image at the red band with un-scanned gaps. (**b**) Imputation using the linear interpolation. (**c**) Imputation using the global-average

359   correction. (**d**) Imputation using the adaptive-average correction. Red boxes highlight the differences of missing-value pixels

360   imputation by different methods (linear interpolation, global-average correction and adaptive-average correction).

361

362 To quantitatively assess the effectiveness of STAIR in the missing-value pixels imputation process, we

363 also used Landsat 5/8 stripe-free images to simulate Landsat 7 SLC-off images, which enabled us to

364 validate the gap filling quality with ground truth values. We compared STAIR to a GNSPI, a gap-filling

365 approach based on neighborhood similar pixel interpolation and has been demonstrated to be able to

366 accurately recover the missing pixels and reduce the striping phenomenon. The two methods were applied

367 to fill the simulated SLC-off images of one heterogeneous (Figure S3a-c) area and one homogeneous area

368 (Figure S34-c). Comparison results showed that STAIR achieved comparable or improved results across

369 six bands for the two test area when compared to GNSPI (Figure S3e-f, Figure S4e-f, Table S1, Table

370 S2). Details of the comparison can be found in Text S1 of Supplementary Materials.

371

372 Taken together, our missing-pixel imputation algorithm is able to fill the cloudy pixels or un-scanned

373 gaps in SLC-off images with high quality. Moreover, the imputation algorithm is robust to nonlinear

374 changing trends and different changing patterns of heterogeneous pixels. Although used as one stage in

375 our Landsat-MODIS fusion pipeline in this work, we highlight that our missing-pixel imputation

376 algorithm can also be used as a stand-alone tool as a generic gap-filling methods for various image

377 analysis and applications.

378 ## 3.2 Fused Landsat-MODIS surface reflectance

379 To quantitatively compare the fusion results of our method with that of STARFM, we first held out the

380 selected subset of Landsat images (Section 2.4) as the test dataset. Our method and STARFM were then

381 applied to available Landsat-MODIS pairs to fuse the fine-resolution image on the date of the ground

382 truth image.

383

384   In order to quantify the extent to which the predicted images agree with the actual images, we first

385   calculated the Pearson correlation between the predicted image produced by our method or STARFM and

386   the ground truth image (Figure 4). This evaluation metric, ranging from -1 to +1, assesses the ability of

387   fusion methods on recovering the correct texture patterns of the target image. The method that can

388   accurately predict the texture patterns of the target image would result in a positive correlation value close

389   to +1, while the method that cannot fully capture the correct texture patterns would lead to a smaller or

390   negative correlation value. We compared the correlation values of our method and STARFM in all

391   predictions on the test dataset across six spectral bands and multiple years from 2000 to 2015. As the

392   predicted images of our method never contain missing values while that of STARFM may contain

393   unfilled pixels (e.g., Figure 5), we only considered those pixels that were not missing values in the

394   predicted images of STARFM when calculated the Pearson correlation, which was a preferable evaluation

395   setting for STARFM. In addition, to make the comparison fair for STARFM, we also excluded the

396   predicted pixels that were derived based on missing-value imputation by STAIR when calculated the

397   correlation, which ensures that the comparison only assess the fusion performance and STAIR will not

398   take advantage of the imputation step when compared to STARFM (see Supplementary Text 2 for

399   details). Scatter plots in Figure 4 show the pairwise comparison of fusion accuracy (evaluated using

400   Pearson correlation) on the green band between our method and STARFM. We found that our method

401   outperformed STARFM when evaluated using the Pearson correlation. In all the years tested in Figure 4,

402   our method generally achieved significantly higher correlation than STARFM (one-sided Mann–Whitney

403   U test P value < 0.001). These results suggest that our method performed better than STARFM on

404   recovering the correct texture patterns in the predicted images. For example, Figure 5 shows the predicted

405   images of our method and STARFM on different dates and different spectral bands. The fine-resolution

406   image predicted by our method successfully captured the correct texture patterns in fine resolution and

407   produced a more realistic fused image than STARFM (Figure 5, left column). In addition, our missing-

408   value pixels imputation algorithm enables the generation of a more unified fused image, while the fused

409   image of STARFM contains visually-noticeable gaps in a SLC-off image (Figure 5, middle column).

410 Moreover, by taking the full advantage of time series of satellite data, our method is able to fill missing-

411 value pixels due to cloud using all available information in the time series, while STARFM is limited to

412 the information in at most two matching image pairs and may fail to impute all cloud-contaminated pixels

413 (e.g, Figure 5, right column). We also observed similar results in other spectral bands. Note that we did

414 not show the result for years 2001, 2004, 2008 and 2012, as there was no Landsat image available in these

415 years that satisfy our criterion of the image selection as test data (Section 2.4). Results of comparison of

416 six different bands (r, g, b, nir, swir1 and swir2) can be found in Figure S1, which suggested that our

417 method generated fused image with higher correlation across all bands and all years when compared with

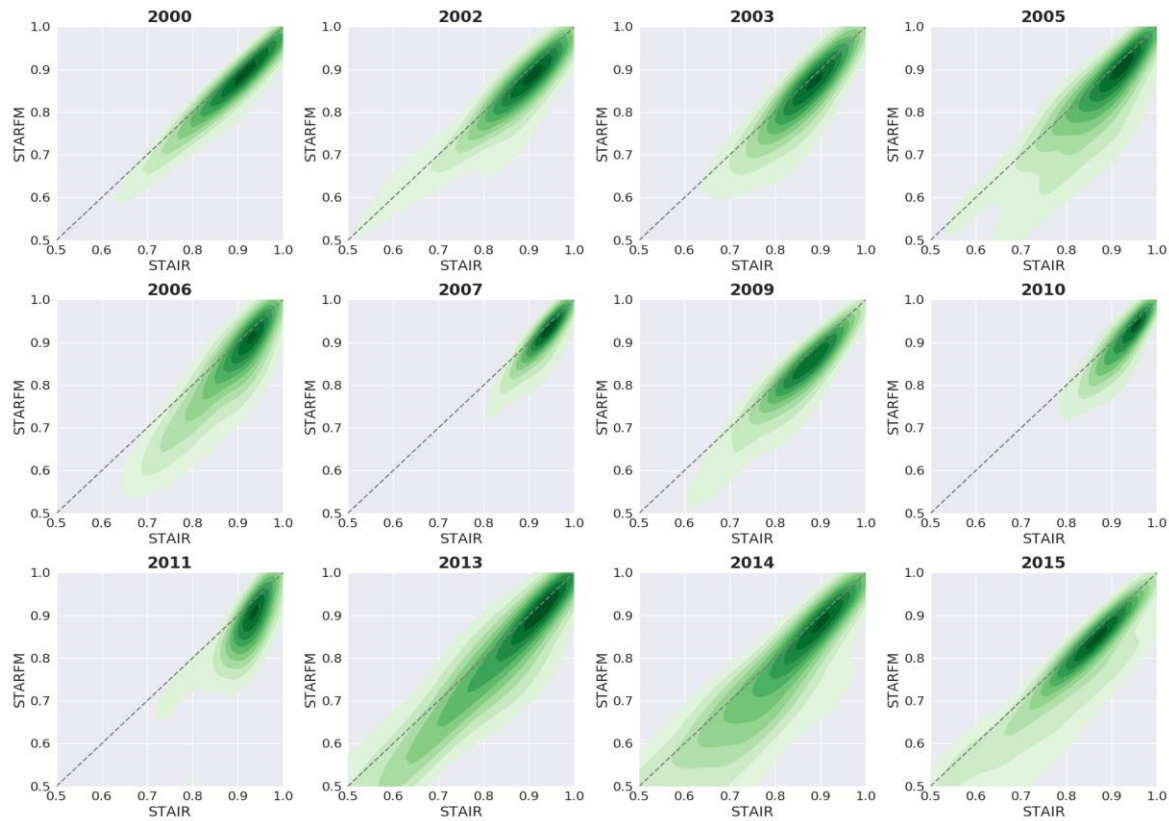418 STARFM.

419

420 **Figure 4. Pairwise comparison of the fusion results of STARFM and STAIR (our method).** The results were evaluated on

421 the Landsat (green band) data spanning multiple years from 2000 to 2015. Pearson correlation was used as the metric to evaluate

422 both methods. The x-axis represents the Pearson correlation between the predicted values by STAIR and the actual values, and

423 the y-axis represents the Pearson correlation between the predicted values by STARFM and the actual values.

424

425   In addition to recovering the realistic texture patterns, it is also desired for a fusion method to predict

426   surface reflectance values at the correct magnitude. To this end, we further evaluate the performance of

427   our method using the root-mean-square error (RMSE) metric (Figure 6). This metric quantifies the

428   absolute difference of surface reflectance values between a predicted image and the ground truth image.

429   Similarly, we computed the RMSE values of our method and STARFM in all predictions on the test

430   dataset across six spectral bands and multiple years. Again, we excluded the pixels with missing value in

431   the predicted images of STARFM when computing the RMSE, even if values of these pixels were not

432   missing in our predicted images, providing a preferable evaluation setting for STARFM. Histogram plots

433   in Figure 6 show the distribution of RMSE values of our method and STARFM when evaluated on the

434   green band. The RMSE values of our method were distributed with significantly smaller mean values than

435   that of STARFM (one-sided Mann–Whitney U test P value < 0.001). Our method also outperformed

436   STARFM for other bands (Figure S2). These results suggest the predicted images produced by our

437   method were more close to the actual images.

438

439   Accurate prediction of the magnitude of pixel values in the fused images greatly benefits the downstream

440   analysis that relies on accurate reflectance values. For example, normalized difference vegetation index

441   (NDVI), which is calculated based on the surface reflectance values of red and nir bands, are often used to

442   monitor crop growth. Our accurate prediction of surface reflectance at red and nir bands would lead to a

443   more accurate NDVI, and thus enable a more reliable crop monitoring. As an example, we built daily

444   NDVI images at 30m resolution for a subarea in the north of Champaign county. Figure 7 shows a subset

445   of these images for every ten days from April 10 to October 20 in 2010. The changing trends of NDVI at

446   field level can be easily tracked in the NDVI images. We further selected two regions with different types

447   of land cover in this area and computed their NDVI time series (Figure 8). The different temporal patterns

448   of the two regions can be easily identified from the fitted curves of the fusion results of our method.

449   These examples showed the potential of our method for being used in downstream applications, such as
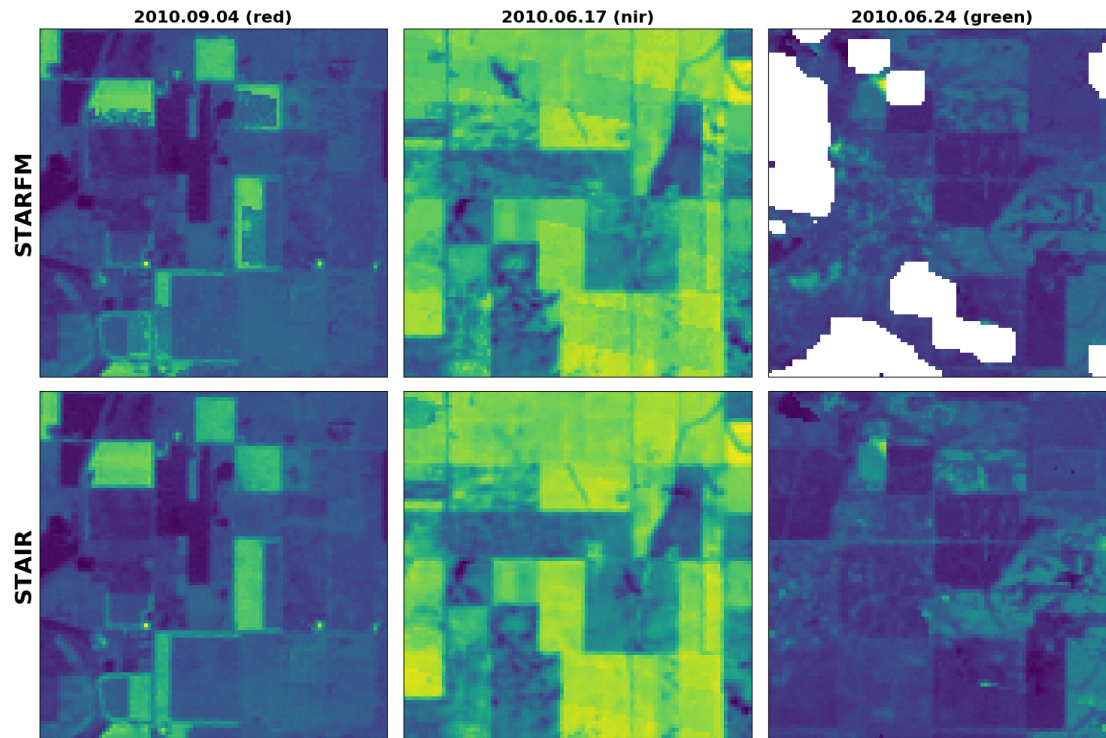
450   quantitative analysis of crop growth.

451

Figure 5. Example of fused images on different dates and different spectral bands. Upper row: STARFM. Lower row:
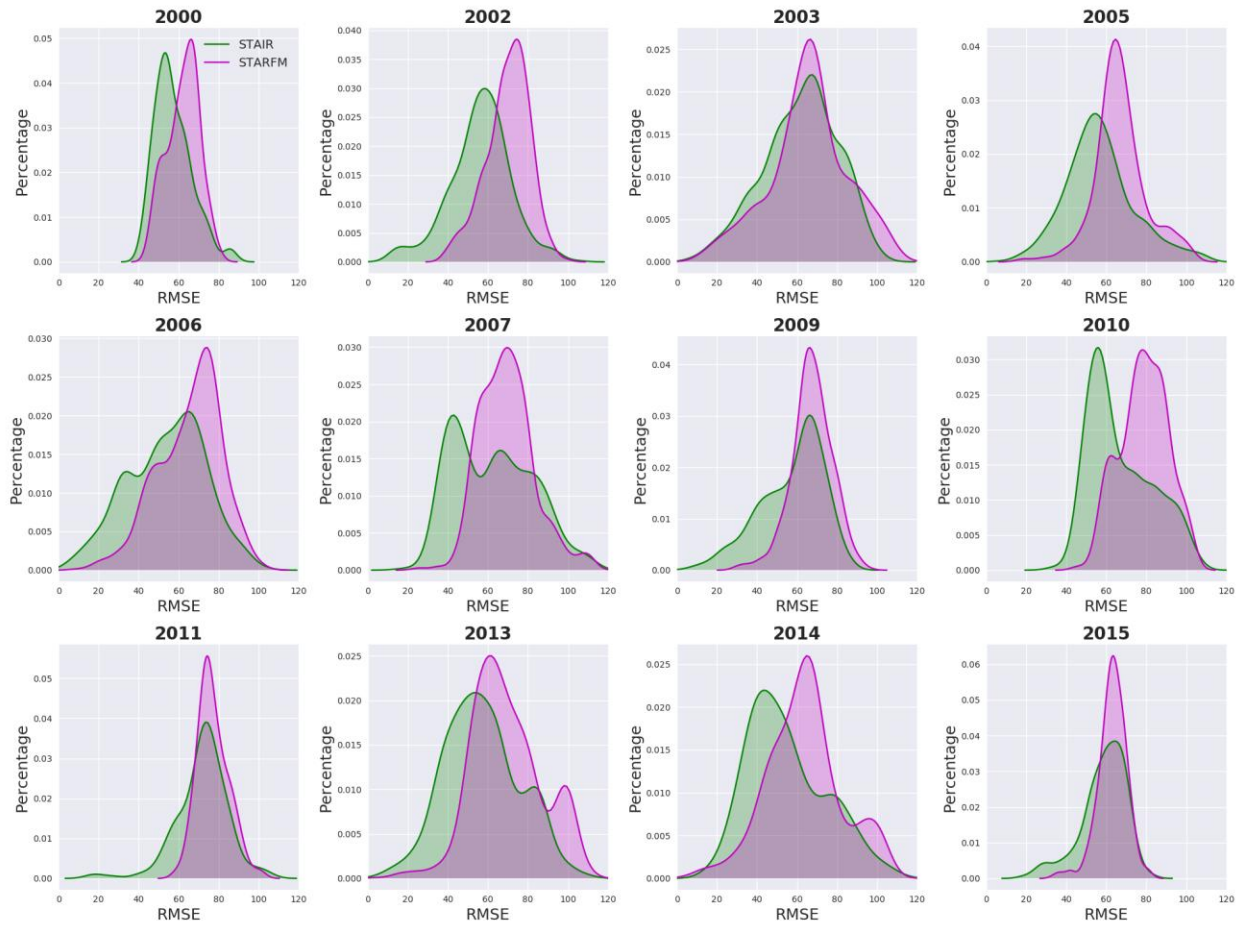
STAIR (our method).

454

455

Figure 6. Histogram of RMSE of STARFM and STAIR (our method). The results were evaluated on the Landsat (green
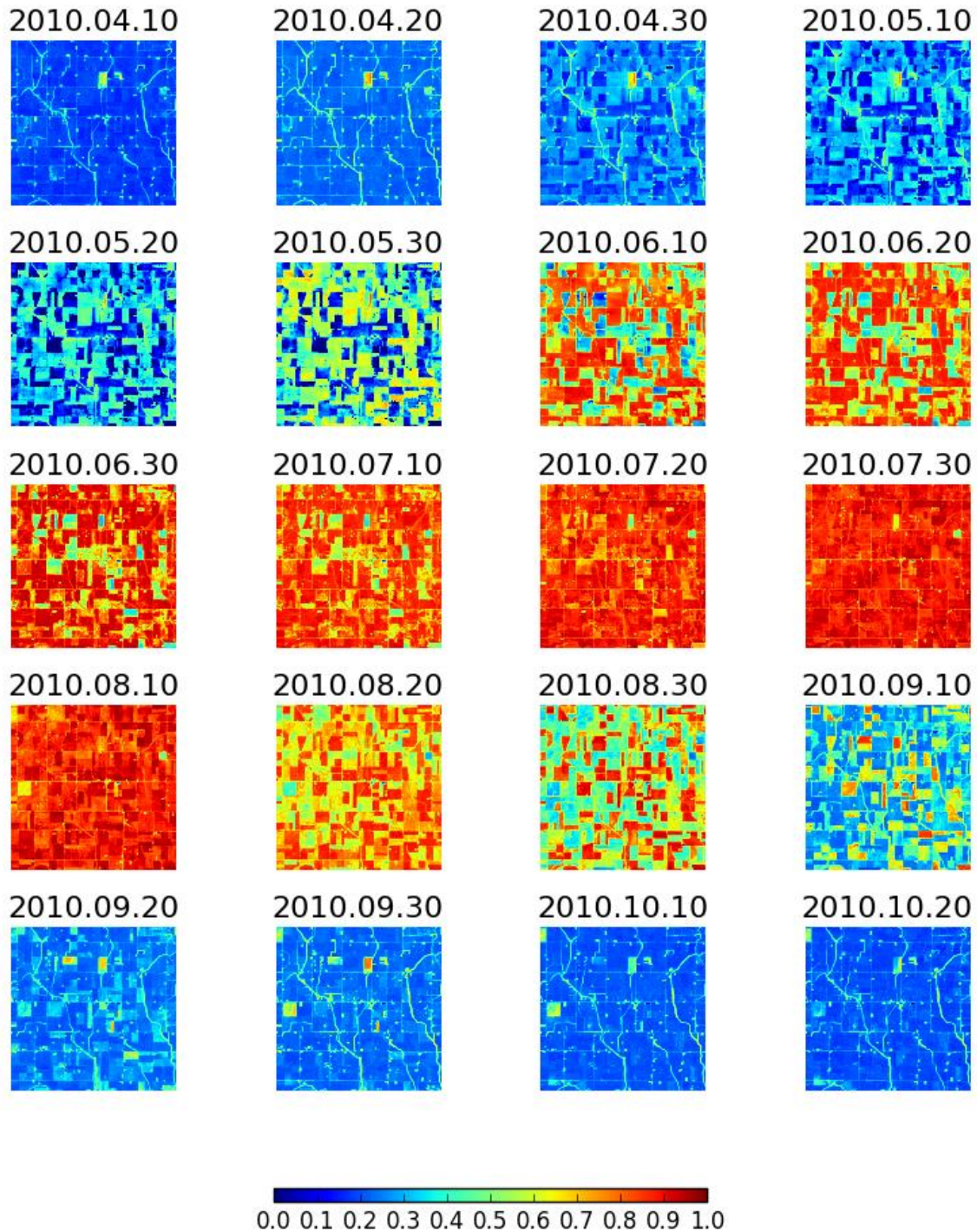
band) data spanning multiple years from 2000 to 2015.

456

457

458

459

460

461

Figure 7. NDVI time series derived from fused images. A subarea (north of Champaign County, IL) was selected to compute the NDVI values from April 10, 2010, to October 20, 2010, covering the growing season.
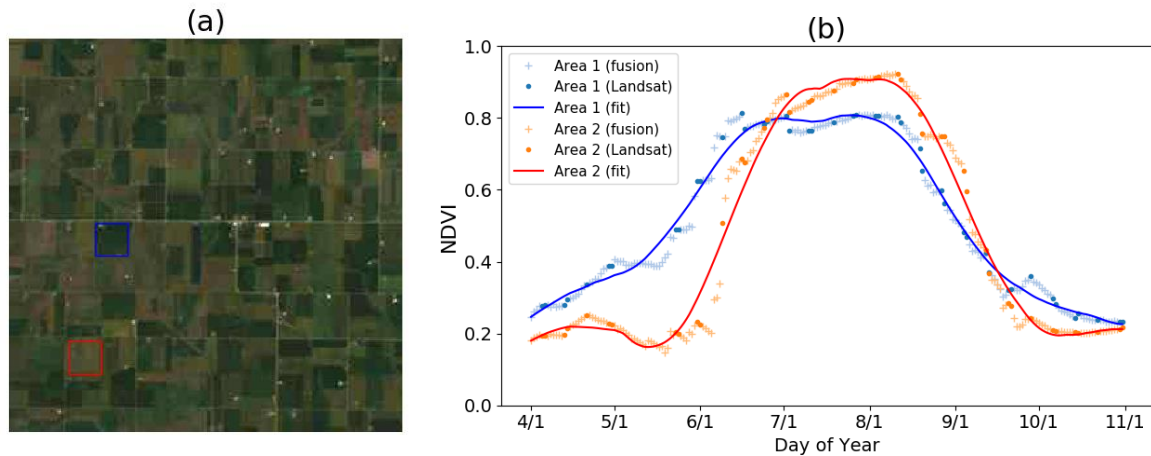
465

Figure 8. Temporal patterns of NDVI. (a) Two regions with different land covers in a subarea (North of Champaign County, IL), Area 1 and Area 2 (highlighted by blue and red boxes, respectively), were selected to compute their NDVI time series. (b) NDVI values derived from Landsat data and fusion data, and fitted curves for Area 1 and Area 2. Fitted curves were built using the Savitzky Golay algorithm.

4. Discussion

4.1 Advancements over previous approaches

Compared to previous methods (e.g., STARFM) for Landsat-MODIS fusion, our method (STAIR) has achieved several advances. The most significant improvements of our method are its flexibility and convenience of usage. Previous Landsat-MODIS fusion methods, including STARFM and ESTARFM, require users to manually select the most proper input Landsat-MODIS pairs (at most two pairs) and specified them as the input of the fusion method. In contrast, when using our method, users only need to simply provide our algorithm with time series of all the available Landsat and MODIS data and do not have to identify pairs for fusion. Our method will then integrate the information in the multiple input pairs, automatically decide which pairs are the most informative, and fuse a fine-resolution image for each of the target dates. The whole pipeline of our method minimizes the human effort and frees users from the tedious process of selecting input pairs, thus it serves as an easy-to-use tool in practice.

25

483

484   Secondly, our method also enables our algorithm to take the full advantage of more available information

485   and produce high-quality fusion products. Our method is able to integrate a time series of satellite data

486   that consists of any number of Landsat-MODIS image pairs for data fusion, while the STARFM

487   algorithm can take at most two pairs of Landsat and MODIS images as input. As found in our

488   experiments, STARFM often failed to fuse image pixels if the corresponding area was missing (e.g.,

489   cloud-contaminated or un-scanned) in both of the two input images pairs. In this case, these pixels

490   remained as missing values in the output image of STARFM. In contrast, by integrating more pairs of

491   Landsat-MODIS image pairs, our method can iteratively fuse the missing pixels until all the pixels are

492   filled.

493

494   Thirdly, our method imputes missing pixels with high quality. As the quality and use of satellite images

495   are often hampered by the cloudy pixels, unscanned pixels and systematic bias, the imputation of missing

496   pixels and the value-correction of imputed pixels are pressingly needed for the fusion products. Our

497   adaptive imputation approach fills missing values in an image by first separating different land covers

498   using an automatic semantic segmentation, and then imputing missing pixels using a locally linear

499   interpolation and a neighborhood correction. Compared to previous gap filling methods, such as NSPI

500   (Chen et al., 2011) and GNSPI (Zhu et al., 2012), the imputation method we developed here can perform

501   the missing-value pixel imputation for all images in the time series at a time. Our method does not require

502   users to specify the reference image, and it will automatically choose reference images that can be used to

503   impute the missing-value pixels. This method can hence be better integrated with our fully-automated

504   fusion algorithm. The high-quality missing-pixel imputation not only facilitates our fusion process, but

505   also serves as a useful stand-alone tool for other remote sensing based analysis.

506

507   Taken together, the results of experiments presented in this work have demonstrated the superior ability

508   of our fusion method, which simultaneously better captures the texture patterns and more accurately

509   predicts the surface reflectance values that are closer to the ground truth. Our method thus provides a

510   high-quality synthetic product with both high spatial and high temporal resolutions, which will benefit

511   other remote sensing based analysis.

512   ## 4.2 A generic framework of fusing multiple-sources of satellite data

513   Our fusion method is a generic framework in that it can fuse multiple types of satellite data. Although in

514   this work we showcase the application of our method on integrating two types of satellite imageries

515   (Landsat and MODIS), the method can be easily extended to further incorporate other types of data,

516   which is a future direction of this work. For example, we can first fuse Landsat and MODIS images as

517   described in this work, and then fuse this synthetic product with a third satellite imagery, e.g., Sentinel-2,

518   using another run of our method (note: cross geo-registration and spectral alignments between different

519   data are required before this step, which is discussed below). This iterative integration approach of

520   multiple satellite imageries only requires little modification to the described method in this work, and a

521   prototype is ongoing.

522

523   Meanwhile, multiple options exist of involving more data sources in our fusion approach. For coarse-

524   resolution satellite data, besides MODIS, VIIRS on board of NASA Suomi-NPP satellite (Cao et al.,

525   2013) and ESA's Sentinel-3 (Donlon et al., 2012) provides a continuation of the MODIS mission. ESA's

526   Proba-V also provides daily and coarse-resolution data (300m-1km) for the similar spectral bands as

527   Landsat (Dierckx et al., 2014). The new generations of geostationary satellite data, e.g. JAXA's

528   Himawari-9 (Bessho et al., 2016), ESA's MSG (Schmetz et al., 2014) and the newly launched NOAA's

529   GOES-R (Schmit et al., 2005), provides sub-hourly satellite data with much more spectral bands similar

530   to MODIS and Landsat (which were not the case for prior generations of geostationary weather satellites).

531   For medium-/high-resolution satellite data, the Landsat mission will continue with expected Landsat 9

532   planned to launch on 2020 (Loveland & Dwyer, 2012); and availability of Sentinel-2 A and B will further

533   increase the sample frequency (though their sampling frequency varies in geographical locations) and also

534 increase the spatial resolution down to 10 m for some spectral bands (Drusch et al., 2012). Though more

535 availability of medium-/high-resolution satellite data can reduce sampling gaps for tropical or cloudy

536 regions, we do not expect that daily sampling at 10-30m resolution becomes possible in the short-term for

537 public satellite data, which justifies the continuous use of the fusion approaches to get daily and high-

538 resolution images.

539

540 Given all these opportunities, cautions need to be given for the best practices before the fusion methods

541 are implemented for different sources of satellite data. Major considerations include unifying spectral

542 responses, cross geo-registration, and sun-sensor angle correction. First, each satellite dataset may have a

543 different coverage of spectral ranges for a certain band, and properly unifying the spectral responses is a

544 prerequisite before data fusion. For example, MODIS's NIR band cover 841-876 nm, Landsat 7 ETM+

545 covers 770-900 nm, and Landsat 8 OLI covers 851-879 nm. Previous studies indicate that surface

546 reflectance values of MODIS and Landsat ETM+ are generally consistent (Masek et al., 2006; Claverie et

547 al., 2015), and spectral responses between ETM+ and OLI are also close enough to be used as a consistent

548 product (Flood 2014; Li et al., 2013). In the current study, we followed the previous approach of

549 STARFM (Gao et al., 2006) to assume that the spectral responses for MODIS, ETM+ and OLI are

550 sufficiently close to be used for fusion purpose, though we fully acknowledge the biases in the cross-

551 sensor spectral responses exist. When fusing Sentinel-2 or other optical satellite data together with

552 MODIS and Landsat, a unification of cross-sensor spectral responses has to be conducted before fusion

553 (Claverie et al., 2016). Second, cross geo-registration is critical for successful data fusion. Each satellite

554 dataset has their own inherent uncertain of geolocations, and geolocation errors also exist between

555 sensors, all of which can propagate to the fusion process. This issue becomes more notable when fusing

556 two high-resolution satellite data (Storey 2016). For example, it has been found that Landsat and Sentinel-

557 2 are currently misaligned by several 10 m pixels (Storey 2016). Robust and automated methods for

558 cross-geo-registration are required for any scaled fusion production, for example, through image

559 texture/spectral matching (Yan et al., 2016). Third, sun-sensor angle correction also requires careful

560 attention and preprocessing. Most high-/medium resolution satellite data are captured from the nadir view

561 (Roy et al., 2016); however, coarse-resolution satellite data usually can have a large variation of viewing

562 angle within a scene due to the large swath coverage. Correcting the BRDF effects for coarse-resolution

563 data is thus necessary before any fusion work (Gao et al., 2006).

564

565 ## 4.3 Assessment of computation efficiency and scalability for continental-scale

566 operational production

567 Our method was implemented in a way such that most of the main steps (such as the imputation of

568 missing-value pixels and the interpolation of Landsat-MODIS) were computed using highly optimized

569 matrix operations, which makes the whole pipeline more computationally efficient than that implemented

570 with pixel-wise operations. In addition, our method can be easily run in a parallel manner. For example,

571 on our local benchmarks, we split the Champaign County into 20 subareas with roughly equal size. We

572 then applied our method on each of the subareas simultaneously, and subsequently merge the fusion

573 results of each subarea to obtain a full fused image of Champaign County. Our method took less than 30

574 minutes on a 32-core machine equipped with 64G memory to generate a product containing daily, 30m

575 resolution images for Champaign County, covering the growing season of one year. This "divide and

576 conquer" framework also enabled us to handle satellite data for larger areas at a continental-scale for

577 operational production, with an only reasonable requirement of running time and memory usage. Our

578 method is currently implemented in Python and runs on CPUs only. In the future, we plan to translate it

579 into a C/C++ version and parallelize it onto GPUs, which would further provide orders of magnitude of

580 acceleration.

581

4.4 Prospects for future satellite earth observations

Though we realize the emergence and increasing availability of satellite datasets with both high resolutions in space and time from the commercial world (e.g. Planet Labs), three major reasons convince us that the fusion methods will continue to play the critical roles in the future satellite earth observation. First, previous and current fusion methods all rely on public and free satellite data, thus the overall cost of generating fused data with both high spatial resolution and high frequency only requires computation cost, which is low for any scaled production. While the industry is still in the phase of promising a daily high-resolution product for the near future, the fusion method has already generated such a daily product as we have described above. Second, scientific values of the fusion data are usually much higher and have gone through more rigorous preprocessing. Specifically, the promised daily image products from the industry usually have limited spectral bands (e.g. visible), and radiance calibration and atmospheric corrections are either lacking or less strictly done, which significantly limits the scientific and more rigorous uses of those satellite images. While the fusion products are based on NASA and ESA's satellite data that has gone through strict scientific preprocessing and can be directly used in both applications and scientific explorations. Third, our fusion method allows us to take the advantages of the rich archives of historical NASA satellite data, and enables to generate historical daily, 30m, cloud-/gap-free images for any places in the planet for the past decades (e.g. since 2000 when MODIS data became available); and the fusion method is the only way to make this happen. These historical fusion data open up opportunities to detect past changes and long-term trends for terrestrial ecosystems and urban areas, and most importantly they allow the possibility to combine with other historical data sources (e.g. statistical survey, observation network) to train advanced models and machine learning algorithms to infer societally meaningful information (e.g. crop yield, economic indicators). With more availability of the public satellite data from NASA, ESA and international community, the fusion products integrating multiple sources would continuously be improved.

607   **List of Acronyms:**

608   BRDF: bidirectional reflectance distribution function.

609   ESA: European Space Agency

610   GOES: Geo-stationary Operational Environmental Satellite

611   JAXA: Japan Aerospace Exploration Agency

612   MODIS: Moderate Resolution Imaging Spectroradiometer

613   MSG: Meteosat Second Generation

614   NASA: National Aeronautics and Space Administration

615   NDVI: Normalized Difference Vegetation Index

616   NOAA: National Oceanic and Atmospheric Administration

617   STAIR: SaTellite dAta IntegRation (name of the presented method)

618   VIIRS: Visible Infrared Imaging Radiometer Suite

619

629

630

631

References

Arvidson, T., Goward, S., Gasch, J., & Williams, D. (2006, 10). Landsat-7 Long-Term Acquisition Plan. *Photogrammetric Engineering & Remote Sensing, 72*(10), 1137-1146. doi:10.14358/pers.72.10.1137

Bessho, K., Date, K., Hayashi, M., Ikeda, A., Imai, T., Inoue, H., . . . Yoshida, R. (2016). An Introduction to Himawari-8/9— Japan's New-Generation Geostationary Meteorological Satellites. *Journal of the Meteorological Society of Japan. Ser. II, 94*(2), 151-183. doi:10.2151/jmsj.2016-009

Cammalleri, C., Anderson, M., Gao, F., Hain, C., & Kustas, W. (2014, 03). Mapping daily evapotranspiration at field scales over rainfed and irrigated agricultural areas using remote sensing data fusion. *Agricultural and Forest Meteorology, 186*, 1-11. doi:10.1016/j.agrformet.2013.11.001

Cao, C., Xiong, J., Blonski, S., Liu, Q., Uprety, S., Shao, X., . . . Weng, F. (2013, 10). Suomi NPP VIIRS sensor data record verification, validation, and long-term performance monitoring. *Journal of Geophysical Research: Atmospheres, 118*(20). doi:10.1002/2013jd020418

Chen, J., Zhu, X., Vogelmann, J. E., Gao, F., & Jin, S. (2011). A simple and effective method for filling gaps in Landsat ETM+ SLC-off images. *Remote Sensing of Environment*, 115(4), 1053-1064.

Claverie, M., Vermote, E. F., Franch, B., & Masek, J. G. (2015, 11). Evaluation of the Landsat-5 TM and Landsat-7 ETM surface reflectance products. *Remote Sensing of Environment, 169*, 390-403. doi:10.1016/j.rse.2015.08.030

Denis, G., Claverie, A., Pasco, X., Darnis, J., Maupeou, B. D., Lafaye, M., & Morel, E. (2017, 08). Towards disruptions in Earth observation? New Earth Observation systems and markets evolution: Possible scenarios and impacts. *Acta Astronautica, 137*, 415-433. doi:10.1016/j.actaastro.2017.04.034

Dierckx, W., Sterckx, S., Benhadj, I., Livens, S., Duhoux, G., Achteren, T. V., . . . Saint, G. (2014, 03). PROBA-V mission for global vegetation monitoring: Standard products and image quality. *International Journal of Remote Sensing, 35*(7), 2589-2614. doi:10.1080/01431161.2014.883097

657 Donlon, C., Berruti, B., Buongiorno, A., Ferreira, M. H., Féménias, P., Frerick, J., ... & Nieke, J. (2012).

658   The global monitoring for environment and security (GMES) sentinel-3 mission. *Remote Sensing*

659   *of Environment*, *120*, 37-57.

660 Drusch, M., Bello, U. D., Carlier, S., Colin, O., Fernandez, V., Gascon, F., . . . Bargellini, P. (2012, 05).

661   Sentinel-2: ESA's Optical High-Resolution Mission for GMES Operational Services. *Remote*

662   *Sensing of Environment, 120*, 25-36. doi:10.1016/j.rse.2011.11.026

663 Flood, N. (2014, 08). Continuity of Reflectance Data between Landsat-7 ETM and Landsat-8 OLI, for

664   Both Top-of-Atmosphere and Surface Reflectance: A Study in the Australian Landscape. *Remote*

665   *Sensing, 6*(9), 7952-7970. doi:10.3390/rs6097952

666 Gao, F., Hilker, T., Zhu, X., Anderson, M., Masek, J., Wang, P., & Yang, Y. (2015, 09). Fusing Landsat

667   and MODIS Data for Vegetation Monitoring. *IEEE Geoscience and Remote Sensing Magazine,*

668   *3*(3), 47-60. doi:10.1109/mgrs.2015.2434351

669 Gao, F., Masek, J., Schwaller, M., & Hall, F. (2006, 08). On the blending of the Landsat and MODIS

670   surface reflectance: Predicting daily Landsat surface reflectance. *IEEE Transactions on*

671   *Geoscience and Remote Sensing, 44*(8), 2207-2218. doi:10.1109/tgrs.2006.872081

672 Gao, F., Anderson, M. C., Zhang, X., Yang, Z., Alfieri, J. G., Kustas, W. P., . . . Prueger, J. H. (2017, 01).

673   Toward mapping crop progress at field scales through fusion of Landsat and MODIS imagery.

674   *Remote Sensing of Environment, 188*, 9-25. doi:10.1016/j.rse.2016.11.004

675 Gaulton, R., Hilker, T., Wulder, M. A., Coops, N. C., & Stenhouse, G. (2011, 02). Characterizing stand-

676   replacing disturbance in western Alberta grizzly bear habitat, using a satellite-derived high

677   temporal and spatial resolution change sequence. *Forest Ecology and Management, 261*(4), 865-

678   877. doi:10.1016/j.foreco.2010.12.020

679 Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., . . .

680   Townshend, J. R. (2013, 11). High-Resolution Global Maps of 21st-Century Forest Cover Change.

681   *Science, 342*(6160), 850-853. doi:10.1126/science.1244693

682 Henson, R. (2008). *Satellite observations to benefit science and society: Recommended missions for the*

683   *next decade*. National Academies Press.

684   Hilker, T., Wulder, M. A., Coops, N. C., Linke, J., McDermid, G., Masek, J. G., ... & White, J. C. (2009).

685   A new data fusion model for high spatial-and temporal-resolution mapping of forest disturbance

686   based on Landsat and MODIS. *Remote Sensing of Environment, 113*(8), 1613-1627.

687   Hossain, F. (2015, 10). Data for All: Using Satellite Observations for Social Good. *Eos, 96*.

688   doi:10.1029/2015eo037319

689   Imhoff, M. L., Zhang, P., Wolfe, R. E., & Bounoua, L. (2010, 03). Remote sensing of the urban heat

690   island effect across biomes in the continental USA. *Remote Sensing of Environment, 114*(3), 504-

691   513. doi:10.1016/j.rse.2009.10.008

692   Jean, N., Burke, M., Xie, M., Davis, W. M., Lobell, D. B., & Ermon, S. (2016, 08). Combining satellite

693   imagery and machine learning to predict poverty. *Science, 353*(6301), 790-794.

694   doi:10.1126/science.aaf7894

695   Ju, J., & Roy, D. P. (2008, 03). The availability of cloud-free Landsat ETM data over the conterminous

696   United States and globally. *Remote Sensing of Environment, 112*(3), 1196-1211.

697   doi:10.1016/j.rse.2007.08.011

698   Li, P., Jiang, L., & Feng, Z. (2013, 12). Cross-Comparison of Vegetation Indices Derived from Landsat-7

699   Enhanced Thematic Mapper Plus (ETM) and Landsat-8 Operational Land Imager (OLI) Sensors.

700   *Remote Sensing, 6*(1), 310-329. doi:10.3390/rs6010310

701   Lloyd, S. (1982). Least squares quantization in PCM. *IEEE transactions on information theory*, *28*(2),

702   129-137.

703   Loveland, T. R., & Dwyer, J. L. (2012, 07). Landsat: Building a strong future. *Remote Sensing of

704   Environment, 122*, 22-29. doi:10.1016/j.rse.2011.09.022

705   Masek, J., Vermote, E., Saleous, N., Wolfe, R., Hall, F., Huemmrich, K., . . . Lim, T. (2006, 01). A

706   Landsat Surface Reflectance Dataset for North America, 1990–2000. *IEEE Geoscience and

707   Remote Sensing Letters, 3*(1), 68-72. doi:10.1109/lgrs.2005.857030

708   Mulla, D. J. (2013, 04). Twenty five years of remote sensing in precision agriculture: Key advances and

709   remaining knowledge gaps. *Biosystems Engineering, 114*(4), 358-371.

710   doi:10.1016/j.biosystemseng.2012.08.009

711   Otsu, N. (1979, 01). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on*

712   *Systems, Man, and Cybernetics, 9*(1), 62-66. doi:10.1109/tsmc.1979.4310076

713   Portmann, F. T., Siebert, S., & Döll, P. (2010, 03). MIRCA2000-Global monthly irrigated and rainfed

714   crop areas around the year 2000: A new high-resolution data set for agricultural and hydrological

715   modeling. *Global Biogeochemical Cycles, 24*(1). doi:10.1029/2008gb003435

716   Pringle, M. J., Schmidt, M., & Muir, J. S. (2009). Geostatistical interpolation of SLC-off Landsat ETM+

717   images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(6), 654-664.

718   Roy, D. P., M. A. Wulder, T. R. Loveland, W. C.E., R. G. Allen, M. C. Anderson, D. Helder, J. R. Irons,

719   D. M. Johnson, R. Kennedy, T. A. Scambos, C. B. Schaaf, J. R. Schott, Y. Sheng, E. F. Vermote,

720   A. S. Belward, R. Bindschadler, W. B. Cohen, F. Gao, J. D. Hipple, P. Hostert, J. Huntington, C.

721   O. Justice, A. Kilic, V. Kovalskyy, Z. P. Lee, L. Lymburner, J. G. Masek, J. McCorkel, Y. Shuai,

722   R. Trezza, J. Vogelmann, R. H. Wynne, and Z. Zhu. 2014. Landsat-8: Science and product vision

723   for terrestrial global change research. Remote Sens. Environ. 145: 154–172.

724   Roy, D., Zhang, H., Ju, J., Gomez-Dans, J., Lewis, P., Schaaf, C., . . . Kovalskyy, V. (2016, 04). A

725   general method to normalize Landsat reflectance data to nadir BRDF adjusted reflectance. *Remote*

726   *Sensing of Environment, 176*, 255-271. doi:10.1016/j.rse.2016.01.023

727   Schmetz, J., Pili, P., Tjemkes, S., Just, D., Kerkmann, J., Rota, S., & Ratier, A. (2002, 07). Supplement to

728   An Introduction to Meteosat Second Generation (MSG). *Bulletin of the American Meteorological*

729   *Society, 83*(7), 991-991. doi:10.1175/bams-83-7-schmetz-1

730   Schmit, T. J., Gunshor, M. M., Menzel, W. P., Gurka, J. J., Li, J., & Bachmeier, A. S. (2005, 08).

731   Introducing The Next-Generation Advanced Baseline Imager On Goes-R. *Bulletin of the American*

732   *Meteorological Society, 86*(8), 1079-1096. doi:10.1175/bams-86-8-1079

733   Stone, B., & Rodgers, M. O. (2001, 06). Urban Form and Thermal Efficiency:How the Design of Cities

734   Influences the Urban Heat Island Effect. *Journal of the American Planning Association, 67*(2),

735 186-198. doi:10.1080/01944360108976228

736 Storey, J. 2016. In Proceedings of the Landsat Science Team Meeting, Virginia Tech, Blacksburg, VA,

737   USA.

738 Schaaf, C. B., F. Gao, A. H. Strahler, W. Lucht, X. Li, T. Tsang, N. C. Strugnell, X. Zhang, Y. Jin, J.

739   Muller, P. Lewis, M. Barnsley, P. Hobson, M. Disney, G. Roberts, M. Dunderdale, C. Doll, P.

740   Robert, B. Hu, S. Liang, J. L. Privette, and D. Roy. 2002. First operational BRDF , albedo nadir

741   reflectance products from MODIS. Remote Sens. Environ. 83: 135–148.

742 Strahler, A. H., Muller, J. P., Lucht, W., Schaaf, C., Tsang, T., Gao, F., ... & Barnsley, M. J. (1999).

743   MODIS BRDF/albedo product: algorithm theoretical basis document version 5.0. MODIS

744   documentation, 23(4), 42-47.

745 Tibshirani, R., Walther, G., & Hastie, T. (2001). Estimating the number of clusters in a data set via the

746   gap statistic. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(2),

747   411-423.

748 Wang, Z., C. B. Schaaf, A. H. Strahler, M. J. Chopping, M. O. Román, Y. Shuai, C. E. Woodcock, D. Y.

749   Hollinger, and D. R. Fitzjarrald. 2014. Evaluation of MODIS albedo product (MCD43A) over

750   grassland, agriculture and forest surface types during dormant and snow-covered periods. Remote

751   Sens. Environ. 140: 60–77.

752 Yan, L., Roy, D., Zhang, H., Li, J., & Huang, H. (2016, 06). An Automated Approach for Sub-Pixel

753   Registration of Landsat-8 Operational Land Imager (OLI) and Sentinel-2 Multi Spectral Instrument

754   (MSI) Imagery. *Remote Sensing, 8*(6), 520. doi:10.3390/rs8060520

755 Yin, G., Mariethoz, G., & McCabe, M. F. (2016). Gap-filling of landsat 7 imagery using the direct

756   sampling method. *Remote Sensing*, *9*(1), 12.

757 Zeng, C., Shen, H., & Zhang, L. (2013). Recovering missing pixels for Landsat ETM+ SLC-off imagery

758   using multi-temporal regression analysis and a regularization method. *Remote Sensing of

759   Environment*, *131*, 182-194.

760 Zhu, X., Chen, J., Gao, F., Chen, X., & Masek, J. G. (2010, 11). An enhanced spatial and temporal

761   adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of*

762   *Environment, 114*(11), 2610-2623. doi:10.1016/j.rse.2010.05.032

763   Zhu, X., Liu, D., & Chen, J. (2012). A new geostatistical approach for filling gaps in Landsat ETM+

764   SLC-off images. *Remote sensing of Environment*, 124, 49-60.

765   Zhu, Z., & Woodcock, C. E. (2012, 03). Object-based cloud and cloud shadow detection in Landsat

766   imagery. *Remote Sensing of Environment, 118*, 83-94. doi:10.1016/j.rse.2011.10.028