A. Project Goals

The goal of this project is to exploit the combined observational capabilities of Landsat and Sentinel-2 to develop the algorithmic, methodological, and computational basis for moderate spatial resolution monitoring of land surface phenology. The specific objectives are:

1. To quantify the timing and magnitude of land surface phenology events (“phenometrics”) at moderate spatial resolution, and

2. To generate gap-filled time series of spectral vegetation indices that characterize the entire seasonal cycle of land surface phenology at fixed time steps.

In support of these goals, we are collaborating with Lars Eklundh at Lund University in Sweden, who developed and continues to refine the TIMESAT software package, which is widely used in the remote sensing community for estimating land surface phenology metrics from remote sensing time series. In addition, over the last 12 months we have collaborated informally with Patrick Hostert and Patrick Griffiths at Humboldt University in Berlin, who are investigating the use of Sentinel 2 for agricultural mapping and monitoring applications. This report covers activities for the second year of this project. As we describe below, we have made significant progress towards our research goals during this period.
B. Year 2 Research Activities

The research focus of this project is on phenology from blended time series of Landsat and Sentinel 2 data. During the reporting period, we focused our activities on data set development and algorithm and refinement using Harmonized Landsat-Sentinel-2 (HLS) data sets that are being generated by NASA for a suite of sites distributed world-wide. Specific activities included:

1. **Continued data set development and algorithm refinement.** Data set development and algorithm refinement focused on two main activities. First, we compiled and evaluated HLS data for a suite of sites in North America, South America, and Europe. Second, we adapted the core algorithm used to create the collection 6 MODIS Land Cover Dynamics product (MCD12Q2) for use with HLS data. A key goal of this effort was to evaluate algorithm performance across a range of land cover types. We present sample results from these activities in Section C (Research Results), below.

2. **Analysis and comparison of phenological metrics estimated from (1) TIMESAT, and (2) a version of the MODIS MCD12Q2 algorithm adapted for use with HLS data.** In addition to adapting the MCD12Q2 algorithm for use at moderate spatial resolution, we worked with our Swedish Colleague (Lars Eklundh) to evaluate results from our approach with those from TIMESAT. The goal of this effort is twofold – to improve overall understanding of algorithmic trade-offs (both computationally and with respect to accuracy), and more generally, to provide the community with guidance regarding best practices for phenological modeling. This effort is ongoing, but results from analyses conducted in Year 2 are consistent with those obtained in the first year of this project, and indicate that both approaches provide similar results, and that most differences are caused by differences in pre-processing (i.e., noise removal, detection of snow, gap-filling, etc.). We currently collaborating with our colleagues in Lund to prepare a manuscript based on this work that documents these results.

3. **Kalman filtering for fusing Landsat and S2A MSI time series.** To investigate alternative approaches to using HLS data, we have developed a Kalman Filter-based algorithm for generating seamless and gap-filled time series of Landsat OLI and S2A MSI reflectances. Kalman filtering is a “state-space” approach to modeling dynamic phenomena that has been widely used in a variety of engineering and science applications, but has received relatively little attention in the remote sensing community. For this work we have developed a fairly simple 1-D model that can be used both with and without a process model (i.e., a prior describing the expected state at each time step). A manuscript based on
this effort is currently in development and is planned for submission in summer 2017.

4. **Compilation of validation data sets to support algorithm testing and refinement.** Validation data sets are essential for assessing the quality of land surface phenology information estimated from our algorithms. Building on efforts from Year 1, we compiled a database of available ground data that includes data from three main sources: in-situ observations from ground observers at LTER sides in the northeastern United States, time series of net ecosystem exchange from eddy covariance towers in the US, and imagery from the PhenoCam network, which provides time series of digital photography from webcams that are configured to monitor vegetation state and seasonal dynamics. We now have in place a fairly extensive data set that spans multiple land cover types, which we are currently using to evaluate algorithm results and test algorithm refinements.

C. **Research Results: Algorithm Assessment**

To evaluate the performance of our algorithm, we have applied it to HLS data at sites in North America, South America, and Europe that include a diverse set of land cover types. Below we present results from a subset of these sites in North America that illustrate the character of data being generated by the algorithm, and that demonstrate the quality of results that it produces.

Figure 1 shows results from an HLS tile in Central Massachusetts that illustrates the effect of data density (and the value of Sentinel 2 data) on algorithm results. The left-hand panel shows the estimated day of year (DOY) corresponding to the start of spring (i.e., green leaf mergence) in 2016. The right-hand panel shows the number of images included in the HLS tile at each 30-meter pixel, and clearly reveals systematic differences in the density of observations across the region that arise from differences in the Landsat and Sentinel 2 tiling systems. Data on the left side of both images come from Landsat 8 only, while data on the right side include both Landsat and Sentinel. The highest density of observations occurs where swaths from adjacent Sentinel orbits overlap. Most critically, the left-hand panel clearly illustrates systematic bias toward earlier DOYs arising from lower data densities. Note that data for the HLS test sites only include imagery from Sentinel 2A, which for the period in question was only acquiring data every 20 days. By 2018, the European Space Agency has committed to acquiring data from Sentinel 2A and 2B, each at 10-day repeat, and so data density will not be an issue in the near future.
Figure 1. Left panel: Start of season day of year from Jan. 1, 2016. Right panel: number of HLS values in the time series at each pixel.

Figure 2. Left: Start of season DOY metric. Right: Digital elevation model. Note that pixels that are not deciduous are masked as “Other” in both images.

In Figure 2, we show results for a sub-area of the HLS tile near Hubbard Brook, New Hampshire, where again, the left-hand panel shows the estimated start of season DOY metric, but the right-hand panel, shows a digital elevation model for the site. These results clearly reveal a topographically induced gradient of roughly 15 days in phenology associated with systematic variation in temperatures related to elevation. Most importantly, these results demonstrate both the realism and the spatial granularity of our results that is not available in coarse spatial resolution LSP products.
In Figure 3, we show results from a much more heterogeneous landscape, centered over the Russel Sage Wildlife Management Area in Louisiana. The management area runs from the upper right quadrant through the center of the image, and is composed of forested wetlands. The surrounding areas include a mix of land use and land cover types including built-up areas, row crops, and livestock agriculture. The image shows the growing season length based on the number of days between the start of season and end of season DOY metrics, and clearly illustrates the signature of different land cover and land use types on land surface phenology.

Finally, in Figure 4 we present results from an agricultural region in Kansas that illustrate multiple dimensions of how time series information can be used to extract useful information related to agricultural land use. Specifically, this figure shows time series for a sample double-crop (winter wheat-soy) and single crop (soy only)
fields in Kansas. In addition to clearly distinguishing between single and double crops, phenological information is useful for distinguishing between different crop types (e.g., soy vs. wheat. vs. maize), and multi-temporal information was used to delineate field boundaries in the crop maps.

**D. Collaborative Interactions with Colleagues in Europe**

We are actively engaged in collaboration with our Swedish colleagues on this project, and in Year 2, we also collaborated with colleagues at Humboldt University in Berlin (Patrick Hostert, Patrick Griffiths). To support this, Friedl, along with post-docs Graesser and Melaas, traveled to Berlin (along with Eklundh, who joined from Sweden) to meet with Hostert and Griffiths on Nov 7 and 8, 2016, and then traveled to Lund to work with Eklundh on November 9 and 10, 2016. In addition, the teams at BU and Lund have quarterly Skype meetings and all three groups are actively engaged in the collaboration. We have 2 manuscripts that are currently in preparation based on our joint work, and we are planning to map out a third paper during our next visit, which is planned for late August of 2017 in Lund.

**E. Project Administration, Personnel, and Budget**

The PI (Friedl) provided oversight of all research activities throughout the reporting period, and devoted 1 month (equivalent) of effort during the summer to this project. In addition, Ms. Radost Stanimirova (Ph.D. student) and Dr. Jordan Graesser (50% FTE post-doc, who started at BU last November) and Dr. Eli Melaas (25% FTE post-doc) all contributed significant effort to project goals during the reporting period.

Because of the gap between Jordan Graesser’s arrival at BU and the departure of Josh Gray (now assistant professor at North Carolina State University), the project was somewhat under-resourced for several months last year, and as a result, the project will be underspent at the end of the current budget period. We currently have plans in place to hire additional staff to catch up on effort and planned deliverables, which will increase our spend rate in Year 3 to compensate.

**F. Ongoing Research and Priorities for Year 3**

Ongoing and planned research for the third year of this project are unchanged from our original work plan. Specifically, we plan to: (1) finalize efforts focused on data set development, (2) finalize our algorithm refinements and evaluation of results at
test sites, (3) continue our collaboration with European colleagues, and (4) devote significant effort to reporting our results via published manuscripts and presentations at national and international scientific meetings.

G. Publications and Presentations During the Reporting Period

During the reporting period, we published one refereed journal article, made one invited seminar presentation, and published one abstract based on a poster presentation at the Fall meeting of the American Geophysical Union, all of which were based on work supported by this grant.

