Staple crops are grown by millions of smallholder farmers yet estimating field-level yields over broad regions can be challenging. Furthermore, agricultural productivity can be impacted by nearby forests and trees. In an agricultural-forest mosaic in Southern Ethiopia, we used remote sensing imagery to identify and differentiate among dominant crops and assess the impact of nearby forest patches on wheat productivity. Using a suite of vegetation indices (VIs) derived from high spatial resolution (5–10 m) satellite imagery as a proxy for wheat productivity, we determined whether VIs were enhanced or suppressed with increasing distance to forest. We found that 5–10 m resolution satellite imagery was sufficient for identifying and differentiating among field boundaries and dominant crops, however, imagery from higher spatial resolution satellites would see increased benefits in accuracy. VIs increased by as much as 5% in areas of the fields within 30 m of forest edges compared to fields further from forests. Our results highlight potential benefits of a landscape approach for enhancing smallholder agricultural productivity in Southern Ethiopia. High spatial resolution imagery is a cost-effective method to map and identify promising landscape approaches in agricultural-forest mosaics dominated by smallholder farms. Thus, a landscape perspective aided by remote-sensing can provide a straightforward and cost-effective way to monitor crop productivity and track changes in agricultural productivity due to forest fragmentation and/or restoration. A landscape approach to achieving food security goals, particularly within the context of climate considerations, and should play a more prominent role in planning forest conservation and restoration.

**Keywords:** forest edge effects, vegetation indices, landscape approach, smallholder agriculture, high spatial resolution imagery, food security, ecosystem services, forest fragmentation
INTRODUCTION

Maize (Zea mays), rice (Oryza sativa), and wheat (Triticum aestivum) are vital to food security, accounting for over half of the world’s dietary caloric supply (FAO, 2018). Global food security has improved during recent decades and the number of people who are undernourished has fallen by more than 200 million people since 1990 (FAO, 2015). However, despite this progress, one in nine people remain undernourished (FAO, 2018). Numerous international, national, and regional initiatives have made food security a priority. For example, the United Nation’s second Sustainable Development Goal (SDG 2) aims to end all hunger and food insecurity by 2030 (UNHCR, 2017). Staple crops are especially important in the least developed countries, where food insecurity is very common (FAO, 2015) and many people rely most heavily on subsistence agriculture for their livelihoods.

Remote sensing is one approach that can help support food security goals by lowering costs and reducing logistical challenges associated with the monitoring of crop yields and other landscape factors that may influence agricultural productivity; yet a problematic data gap exists for smallholder farms. To date, a greater emphasis has been placed on remote sensing of yields at broad spatial extents and coarse spatial resolutions, and often on industrial-scale agriculture in developed countries (Mulla, 2013). For example, yields of staple crops such as maize and wheat are routinely monitored over large areas using coarse or moderate resolution satellite imagery such as MODIS (Moderate Resolution Imaging Spectroradiometer) and Landsat (Doraisswamy et al., 2004; Ren et al., 2008). However, such approaches are unable to differentiate heterogeneous crops in smallholder fields (Grassini et al., 2015).

Almost 85% of the world’s estimated 570 million farms are smaller than 2 ha in size (Lowder et al., 2016). As with any farm, assessing the productivity and health of smallholder farms can help identify problems and improve yields. However, obtaining direct yield measurements from small, fragmented, and highly diverse fields and farms common to smallholders can be particularly challenging. High spatial resolution satellite imagery such as Sentinel-2 (10 m), RapidEye (5 m), and Quickbird (2.5 m), can help fill this data gap for smallholder fields otherwise impossible to analyze with coarser resolution imagery. In addition, high spatial resolution imagery also enables identification of small yet important fine resolution features such as small forest patches and scattered trees which can impact agricultural productivity (Pu and Landry, 2012).

Forests and scattered trees can help improve food security by providing ecosystem services and improving agricultural productivity (Zomer et al., 2009; Bommarco et al., 2013; Reed et al., 2017). For example, when grown alongside staple crops as part of an agroforestry system, Faidherbia albida (F. albida) can improve water availability, soil moisture, and water use efficiency (Laike, 1991; Camargo and Kapos, 1995; Jackson et al., 2000; Payne, 2000; Sida et al., 2018a). This tree species can also enhance soil fertility via nitrogen fixation (Giller, 2001; Kho et al., 2001), organic matter build-up (Gelaw et al., 2014), and nutrient cycling (Cannell et al., 1996). Importantly, forests near field margins also moderate microclimatic conditions within fields (Gehlhausen et al., 2000; Chen et al., 2010), enhancing agricultural yields especially for C3 crops such as wheat (Reed et al., 2017). In certain ecosystems, forests and trees may provide habitat for both pollinators and natural enemies that prey on agricultural pests, thereby increasing yields (Dix et al., 1995; Bianchi et al., 2006; Garibaldi et al., 2013).

The landscape mosaic of smallholder farms interspaced with forest patches throughout Southern Ethiopia represents an ideal location for exploring relationships between forest edges and agricultural productivity. Here, smallholder farms are often composed of a complex mosaic of croplands, pastures, home gardens, and forest patches integrated into a mixed crop-livestock system (Baudron et al., 2017). Wheat is a major staple crop for smallholders in the cooler Ethiopian highlands (Tanner and Hulluka, 1991) and is among the primary cereal crops grown in Ethiopia, following only teff (Eragrostis tef) and maize in terms of total production and cultivated area (Taffesse et al., 2012). Additionally, wheat yield is extremely susceptible to microclimatic effects from nearby forest edges. Within this context, we explore two primary questions regarding the relationship between forests and wheat productivity:

Can high spatial resolution remote sensing be used to distinguish among crops and estimate productivity on smallholder farms?

Using recently available high spatial resolution satellite imagery (RapidEye and Sentinel-2), we aimed to identify and distinguish between wheat and maize fields and measure patterns of agricultural productivity in smallholder farms. RapidEye is a series of commercial satellites capturing 5 m resolution imagery from five multi-spectral sensors. The geographic extent of image capture depends on commercial demand, and therefore, archived historical imagery is limited. In contrast, the European Space Agency Sentinel-2 satellites with 10 m resolution imagery has a free and open data policy and an average global revisit time of 5 days. To distinguish wheat and maize fields, we first use RapidEye and Sentinel-2 imagery to discriminate field boundaries and identify crop types based on seasonal differences due to vegetation phenology and field preparation (including timing of plowing and handling of crop residue).

Does proximity to forest edge impact wheat productivity in smallholder farms?

Second, we also explored whether forest proximity affects sub-field-level wheat productivity using remotely sensed vegetation indices (VIs). A vegetation index is a simple empirical index derived from two (or more) spectral responses, which can be used as a surrogate of vegetation conditions such as plant health and vigor. Agricultural productivity can be estimated in this way because higher productivity and healthier crops are often associated with higher VIs (Hatfield, 1983; Doraiswamy et al., 2004; Ferencz et al., 2004; Burke and Lobell, 2017; Jin et al., 2017). Because certain VIs exhibit higher correlations...
with crop status, we assessed and compared several VIs. Using VIs we examined how agricultural productivity is affected by forest edges. We hypothesized that VI will be highest in sections of the field closest to forests due to potential benefits from forest edges. Understanding this relationship as well as how the configuration of landscape mosaics affects crop productivity is key to improving food security by helping farmers better optimize agricultural productivity at the landscape level.

**MATERIALS AND METHODS**

To answer these questions, we used a suite of field-based assessments and high spatial resolution imagery from the Sentinel-2 and RapidEye satellites. We first classified the landscape into general land cover types and then distinguished two major crop types (wheat and maize) using ground verification. After isolating fields of wheat adjacent to forest patches, we analyzed wheat productivity at different distances from forest patches using a suite of vegetation indices. Next, we explain each of these steps in more detail.

**Study Region**

The 50 km² study region is south of the capital city of Addis Ababa, Ethiopia, within the Arsi-Negele woreda (district) of the Oromia Region and situated 40 km northeast of the city of Awassa (Figure 1). The study region has a prominent forest-agricultural gradient, which provides a large sample of fields of varying proximity to forest. Borded by the state forest of Munesa to the east and the town of Arsi-Negele to the west, forest patches as well as individual trees become larger and more abundant moving eastward (Baudron et al., 2017). The main crops grown in the area are wheat, potato, and maize, often integrated with livestock as part of a mixed farming system (Baudron et al., 2017). The main tree species in the forested areas are *Podocarpus falcatus*, *Eucalyptus sp.*, *Croton sp.*, *Pinus patula*, and *Prunus Africana* (Baudron et al., 2019).

The site is situated within the oceanic climate (Cfb) zone (Kottek et al., 2006). The Degaga weather station (~20 km north from the study region) reports a mean annual rainfall of 1,075 mm (18-year mean) and a mean annual temperature of 15°C (16-year mean) (Halle-Wittenberg University, 2002). A bimodal rainfall pattern produces a shorter rainy season from March to May and a longer rainy season from July to September.
As such, the growing season encompasses two distinct periods impacted by these two rainy seasons. During the shorter rains, land is either fallowed or used to grow potato. During the longer rains, wheat is generally planted following either the fallow or potato crop. Additionally, maize is planted at the beginning of the shorter rains and harvested at the end of the longer rains.

Image Classification
To answer our first research question, we classified high spatial resolution RapidEye imagery into general land cover classes and then used Sentinel-2 imagery to further differentiate crop types within agriculture.

RapidEye 2015 Classification
To classify general land cover, we acquired orthorectified RapidEye 3A imagery (5 m resolution) from January 3rd, 2015. Orthorectification of RapidEye imagery is based on ground control points and digital elevation models, and has an accuracy of 10 m RMSE (root mean square error). The post-harvest dry season (December-February) was selected to enable optimal differentiation between classes using seasonal phenology of vegetation. We performed a maximum likelihood supervised classification to distinguish among three broad land cover classes: (1) forests, (2) grassland (which included pastures, meadows, and floodplains), and (3) agricultural fields and/or bare soil. Around 200 training sites (to train the classification algorithm) were visually delineated from the RapidEye imagery using Google Earth imagery (from January 24th 2014). The minimum mapping unit for forests was 125 m² (i.e., five RapidEye pixels) which helped minimize the effects of misclassification by ensuring single trees and extremely small forest patches were carefully delineated/excluded from the RapidEye classification.

Sentinel 2016 Classification
Orthorectified (surface reflectance based) Sentinel-2 2A imagery (10 m resolution) from December 31st 2016 was used to further separate the agricultural land into two classes, wheat and maize, using ENVI 5.4. Imagery from the post-harvest period (i.e., December–February) was particularly useful in differentiating wheat and maize. From this imagery, we performed an object-based classification to group homogenous pixels into spectrally similar clusters (similar to patches). First, within the areas corresponding to the agricultural areas of the RapidEye imagery, the Sentinel-2 imagery was segmented into clusters of pixels with homogenous characteristics. This procedure used the “edge” and “full lambda schedule” merge algorithms. The levels of edge and merge were set at 50 and 80, respectively, to enable segmentation of clusters approximating average field sizes in the region (~0.3 ha).

Second, classification of clusters was based primarily on their spectral characteristics. The selected image date optimally differentiated the two crops because of the distinct bright signature of straw residue left behind on wheat fields immediately after harvest. In contrast, post-harvest maize fields were differentiable by their darker spectral signature associated with bare soil (because maize crop residues are typically removed and used as livestock feed). Training data for the classification was selected from the Sentinel-2 imagery guided by these spectral distinctions.

Ground Verification
An accuracy assessment was performed on both the RapidEye 2015 classification and the Sentinel-2 2016 classification. Seventy-one randomly selected ground verification points collected via field work in April 2015 were used to assess the RapidEye classification. Additionally, 25 photos of fields with identifiable crop residue were taken in February 2017 and used as ground verification points for the Sentinel-2 classification. We created a confusion matrix to evaluate the producer’s, user’s, and overall accuracy, as well as the Kappa coefficient. User’s accuracy helps identify over-mapping of a class, as user’s accuracy is inversely proportional to errors of commission. Producer’s accuracy is inversely related to errors of omission (or exclusion), indicating a given class was missed and erroneously mapped as another class (Congalton and Green, 2008). The kappa statistic is a measure of the quality of the classification relative to assignment of classes by random chance (Congalton and Green, 2008).

Forest Edge Yield Relationships Using Vegetation Indices
To answer the second question, we conducted a proximity analysis using the results of the classified wheat map and then derived a suite of vegetation indices from Sentinel-2 imagery from October 31st, 2016 just prior to harvest. We evaluated three VIs for comparative purposes: modified soil-adjusted vegetation index 2 (MSAVI2), normalized vegetation index (NDVI), and green chlorophyll vegetation index (GCVI) (detailed in Table 1). NDVI is a simple indicator calculated using the red and infrared bands and was selected due to its widespread use among various scientific and natural resource applications (Xue and Su, 2017). However, as NDVI products have been shown to produce incorrect results when it comes to landscapes with exposed soil, MSAVI2 was included to minimize the impact of bare soil on canopy reflectance in a way that differs from other soil-adjusted indices by quantitatively streamlining the soil brightness correction (Qi et al., 1994). Lastly, we included GCVI, a VI utilizing green reflectance, which is more responsive to leaf chlorophyll than the more oft-used red reflectance. Leaf chlorophyll is a photosynthetic pigment in plants that correlates with nitrogen levels (Shangguan et al., 2000; Cartelat et al., 2005). As a result of this link with plant nutrient deficiency, GCVI is

---

**TABLE 1 | Mathematical formulas and references for the three vegetation indices used in this study.**

<table>
<thead>
<tr>
<th>Vegetation Index</th>
<th>Formula</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDVI (Normalized Difference Vegetation Index)</td>
<td>( \frac{\text{Red} - \text{NIR}}{\text{Red} + \text{NIR}} )</td>
<td>Rouse et al., 1973</td>
</tr>
<tr>
<td>MSAVI2 (Modified Soil-Adjusted Vegetation Index 2)</td>
<td>( \frac{2(\text{NIR} + 1) \sqrt{\text{NIR} + 1} - \text{NIR} - \text{RED}}{\text{NIR} - 2} )</td>
<td>Qi et al., 1994</td>
</tr>
<tr>
<td>GCVI (Green Chlorophyll Index)</td>
<td>( \frac{\text{NIR}}{\text{RED}} - 1 )</td>
<td>Gitelson et al., 2003</td>
</tr>
</tbody>
</table>
TABLE 2 | Raw tally confusion matrix of land cover and accuracies of the RapidEye (January 3rd 2015) classification.

<table>
<thead>
<tr>
<th>Classification data</th>
<th>Agriculture and bare soil</th>
<th>Grassland</th>
<th>Forest</th>
<th>Plantations</th>
<th>Grand total</th>
<th>User's accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture and bare soil</td>
<td>15</td>
<td>1</td>
<td></td>
<td></td>
<td>16</td>
<td>0.94</td>
</tr>
<tr>
<td>Grassland</td>
<td>1</td>
<td>16</td>
<td></td>
<td></td>
<td>17</td>
<td>0.94</td>
</tr>
<tr>
<td>Forest</td>
<td></td>
<td></td>
<td>19</td>
<td>1</td>
<td>20</td>
<td>0.95</td>
</tr>
<tr>
<td>Plantations</td>
<td></td>
<td></td>
<td>18</td>
<td></td>
<td>18</td>
<td>1.00</td>
</tr>
<tr>
<td>Grand total</td>
<td>16</td>
<td>17</td>
<td>19</td>
<td>19</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Producer's accuracy</td>
<td>0.94</td>
<td>0.94</td>
<td>1</td>
<td>1</td>
<td>0.95</td>
<td></td>
</tr>
</tbody>
</table>

Seventy one total verification points were used to assess accuracy for this classification. Overall accuracy is 0.96 while the Kappa statistic is 0.83.

TABLE 3 | Raw tally confusion matrix of land cover and accuracies of the Sentinel-2 (December 30th 2016) classification.

<table>
<thead>
<tr>
<th>Classification data</th>
<th>Wheat</th>
<th>Maize</th>
<th>Grand total</th>
<th>User's accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat</td>
<td>13</td>
<td>2</td>
<td>15</td>
<td>0.87</td>
</tr>
<tr>
<td>Maize</td>
<td>0</td>
<td>10</td>
<td>10</td>
<td>1.00</td>
</tr>
<tr>
<td>Grand total</td>
<td>13</td>
<td>12</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>Producer's accuracy</td>
<td>1.00</td>
<td>0.83</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Twenty five total verification points were used to assess accuracy for this classification. Overall accuracy is 0.91 while the Kappa statistic is 0.83.

Statistical Analysis
To examine patterns of wheat productivity, we compared VIs within sub-sections of wheat fields, where sections of the field were located at different distances from forest. To do so, we first delineated boundaries of all individual wheat fields (using the above classification). From these, potential "treatment" fields were identified as those adjacent to a forest patch (using a minimum forest patch width and length of 20 and 80 m, respectively). From these fields, 40 fields were randomly selected and further partitioned into "treatments" based on distance from forest edge: 10–30, 31–40, 51–70, 71–90 m. Additionally, we left a 10 m gap between the first distance class (10–30 m) and the forest edge to minimize potential "mixed-pixel" problems (i.e., tree crowns overlapping into the first distance class). Lastly, we randomly selected another 25 fields which were more than 200 m from any forest patch to serve as “control” fields which required some important assumptions. A distance of 200 m was chosen because it is far enough away that microclimatic effects from forest edges would be negligible (Davies-Colley et al., 2000), however, ecosystem services such as pollination and pest control could extend more than 200 m from forests (Chacoff et al., 2008; Ricci et al., 2009; Bailey et al., 2014). Further, the general configuration and interspersion of forests in the study area did not allow for creation of a sufficient number of controls at distances any further than 200 m. At the pixel level, we compared VIs among distance classes using a linear model. Additionally, we conducted post-hoc pairwise comparisons using Tukey’s range tests. Each distance class within each field had various replicates (pixels), with a total of 2,431 pixels in treatment fields and 1,611 in control fields. Lastly, to evaluate the underlying assumed relationship between VI and wheat yield, we conducted a simple linear regression between ground verified wheat yields in 11 fields to VIs from 2014 imagery.

RESULTS
High Spatial Resolution Remote Sensing Can Distinguish Wheat and Maize Fields and Estimate Wheat Productivity on Smallholder Farms
The RapidEye image was classified into three broad classes (forest and tree cover, grasslands, and agricultural areas and bare soils) (Figure 1). Overall accuracy of the RapidEye classification was 96%, with a kappa statistic of 0.83 (Table 2). Both the producer's and user's accuracy ranged from 94 to 100%. The two land cover classes of most importance to our analysis, forest and agriculture, had high user's and producer's accuracies.

A subset of RapidEye-derived agriculture class was then classified into wheat and maize using Sentinel-2 imagery (Figure 1). The mapped extent of wheat was 1,806 ha whereas the total extent of maize was 1,835 ha. Overall accuracy, based on 25 ground verification points, was 91% (Table 3). Per class user's accuracies for wheat and maize was 81 and 78, respectively, and producer's accuracy was 100 and 83%, respectively.

Through a simple linear regression analysis we found a moderately strong, positive relationship between NDVI and yield ($R^2 = 0.3675$) (Figure 2).
Closer Proximity to Forest Edges Increases Indices of Wheat Productivity in Smallholder Farms

VI values throughout all fields ranged from 0.55–0.78 for MSAVI, 0.52–0.81 for NDVI, and 1.14–3.25 for GCVI. Generally, VIs decreased with distance from forest edge. Mean VI values were highest at the 10–30 m distance class and generally decreased with increasing distance from the forest (Table 4, Figure 3). Mean VI at 10–30 m was significantly higher than at 31–50 m, 51–70 m, and 71–90 m (Table 4). However, the latter three distance classes were generally not significantly different from one another with a few exceptions (Table 5, Figure 3). Mean VI values of all distance classes were generally higher than control fields located >200 m away from forest edges (Table 4). Only the first distance class (10–30 m) was statistically significantly different from the controls ($p \leq 0.0001$) for all three VIs, while the second distance class was significantly different from the control for GCVI and MSAVI2 (Table 5).

DISCUSSION

Forest restoration is receiving renewed attention globally and particularly in Ethiopia Chazdon, 2008; Pistorius et al., 2017. Thus, clarifying the relationship between forests and crop productivity could aid forest restoration as well as forest conservation planning. By providing guidance for how to optimize landscape features to provide multiple benefits, it is conceivable that multiple stakeholders, from subsistence smallholder farmers to governments planning carbon sequestration programs could benefit. Our research estimated agricultural productivity using a suite of vegetation indices (VIs) derived from remote sensing of wheat and maize fields within a forest mosaic. We found that wheat fields could be well-discriminated from other crop types (i.e., maize) and that forests were associated with higher wheat yields, with some caveats.

Successful Mapping of Crop Types and Productivity Using High Resolution Remote Sensing on Smallholder Farms

We were able to successfully distinguish between two crops using high spatial resolution imagery with classification rules based on seasonal harvest patterns. Overall, user’s, and producer’s accuracies for both the RapidEye and Sentinel-2 classification were all high (exceeding 85% in most cases). Additionally, specific classes of interest (forest, agriculture, and wheat in particular) were mapped with a high degree of accuracy. Arguably, some of the success in achieving high accuracies is attributable to implementation of a classification scheme with few classes (Olofsson et al., 2014). However, we also captured and incorporated land use practices (fallows and crop residues) in a novel way in heterogeneous smallholder farms. The use of post-harvest field phenology for crop identification (light vs. dark spectral signatures for wheat and maize, respectively) uniquely and effectively differentiated between these two crop types.
types. This approach was both simple and cost-effective in that it only required one image from a post-harvest date. However, our approach greatly benefited from familiarity with local farming practices, on-the-ground local knowledge, and fortuitous spectral differences in residual post-harvest agricultural biomass.

Crop identification using remote sensing is often complex, especially in small heterogeneous farms. Crop identification often uses coarse spatial resolution image time series over large fields, often on large commercial farms. For example, a “crop proportion phenology index” was created from a MODIS (250 m) image time series to estimate wheat productivity in China (Pan et al., 2012). A harmonic analysis on time-series data from National Oceanic Atmospheric Administration’s (NOAA) advanced very high resolution radiometer (AVHRR) (1.1 km) was used to estimate maize, soybean and alfalfa in the USA (Jakubauskas et al., 2003) and MODIS-EVI time series was used to identify rice paddies in Japan (Sakamoto et al., 2013). While these studies aid in crop monitoring of large fields, the same geodata and methods are unlikely to be successful on smallholder fields (for example, a 2 ha farm is roughly 200 by 100 m).

Due to its increasing availability, the use of higher spatial resolution imagery for crop identification has expanded. For example, in India, sugarcane, ragi, paddy, and mulberry were distinguished using Quickbird imagery (2.4 m) with high overall accuracies (87–98%) (Senthilnath et al., 2008). Cotton, winter wheat, and rice were identified in Uzbekistan using a combination of imagery from the Satellite Pour l’Observation de la Terre (SPOT) (2.5–5 m) and advanced spaceborne thermal emission and reflection radiometer (ASTER) (15–30 m) satellites with an overall accuracy of 80% (Conrad et al., 2010). With improved access to higher resolution imagery, remote crop identification for smallholder farms is becoming increasingly feasible. Notably, imagery at these higher spatial resolutions is also capable of characterizing detailed off-farm features.

### TABLE 4 | Mean and LS means across all distance classes and control for all three VIs.

<table>
<thead>
<tr>
<th>Distance Class</th>
<th>NDVI Mean</th>
<th>NDVI Least squares mean</th>
<th>GCVI Mean</th>
<th>GCVI Least squares mean</th>
<th>MSAVI2 Mean</th>
<th>MSAVI2 Least squares mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>10–30 m</td>
<td>0.636</td>
<td>0.646</td>
<td>2.413</td>
<td>2.413</td>
<td>0.716</td>
<td>0.716</td>
</tr>
<tr>
<td>31–50 m</td>
<td>0.627</td>
<td>0.627</td>
<td>2.352</td>
<td>2.352</td>
<td>0.709</td>
<td>0.709</td>
</tr>
<tr>
<td>51–70 m</td>
<td>0.623</td>
<td>0.623</td>
<td>2.311</td>
<td>2.311</td>
<td>0.708</td>
<td>0.708</td>
</tr>
<tr>
<td>71–90 m</td>
<td>0.619</td>
<td>0.619</td>
<td>2.294</td>
<td>2.295</td>
<td>0.707</td>
<td>0.707</td>
</tr>
<tr>
<td>Control</td>
<td>0.622</td>
<td>0.622</td>
<td>2.262</td>
<td>2.262</td>
<td>0.703</td>
<td>0.703</td>
</tr>
</tbody>
</table>

**FIGURE 3 |** Plot of least squares mean for NDVI, GCVI, and MSAVI in wheat fields at distances of 10–30, 31–40, 51–70, 71–90 m, and control. Derived from linear model least squares means. Error bars denote standard error. \( n = 40 \) for all distance classes, \( n = 25 \) for control. *\( P \leq 0.05 \), **\( P \leq 0.01 \), ***\( P \leq 0.001 \).
important to smallholder food security and health, including forests and scattered trees (Immitzer et al., 2012), and small water resources (Sawaya et al., 2003).

### Relating Agricultural Yields to Vegetation Indices (VIs)

While this study does not directly estimate yield, we considered VIs to be a useful proxy for productivity due to the abundant literature supporting a strong correlation between the two along with our ground verification in 11 fields. Higher vegetation productivity is often indicative of greater yields as strong linear relationships exist between VIs and crop yield (Hatfield, 1983; Doraiswamy et al., 2004; Ferencz et al., 2004; Burke and Lobell, 2017; Jin et al., 2017). Despite the influence of additional factors on yield (such as temperature, precipitation, soil properties, and fertilizer inputs), this linear relationship is often used to directly estimate yields using simple linear regression. For example, maize and soybean yields were estimated in Iowa using NDVI (at 8 km resolution from the AVHRR), achieving model $R^2$ of 0.78 and 0.86, respectively (Prasad et al., 2006). In Shandong, China, winter wheat yields were also estimated using MODIS-derived NDVI with similar levels of agreement ($R^2 = 0.66–0.87$) (Ren et al., 2008). Successful mapping of yields on broad-scale commercial farms using VIs derived from low-resolution imagery has been common over the past decade.

More recently there have been increasing efforts to map and estimate crop yield using high resolution imagery on smaller farms, oftentimes in developing countries. In Kenya, yield from smallholder farms was estimated from GCVI derived from Terra Bella imagery (1 m resolution) (Burke and Lobell, 2017). Empirical estimation of yields using VIs and ground measurements was compared to a model comparing VI and simulations from APSIM (Agricultural Production Systems Simulator) with reasonable agreement ($R^2 = 0.4$) (Burke and Lobell, 2017). Furthermore, of all the VIs assessed (EVI, NDVI, and GCVI) in the study, GCVI consistently showed the best results in terms of agreement with ground-measured yields (Burke and Lobell, 2017). Similar methods replicated in Zambia and India achieved successful results (Jain et al., 2016; Azzari et al., 2017).

Using a limited dataset from our study area we also tested the relationship between VI and wheat yield, and found a significant relationship between wheat yield and NDVI ($R^2 = 0.3675$, $p = 0.048$). Our methods constitute a reliable way of estimating relative productivity throughout a landscape, but further work would benefit from exploring yield estimates using APSIM simulations calibrated with detailed soil and weather data and/or additional field verification.

### Effects of Proximity to Forests on Wheat Yields

Our proximity analysis detected an increase in VI in field margins situated closest to forests. Regardless of the specific VI used, mean/median VIs were generally higher in areas of the fields nearer to than farther from forests. Considering the range and equation for NDVI and MSAVI2 for example (Figure 3); an increase of 0.01 in VI (half of the difference seen between the first distance class and control) would equate to around a 3–5% increase in productivity. For context, yield increases of 2–14.5% occurred with the addition of fertilizer in two trials in Argentina and the USA (Diaz-Zorita et al., 2004; Freeman et al., 2006).

While trees and forests within agroforestry systems will compete with crops for sunlight, water, and nutrients, they could also provide many positive benefits as well. Trees from agroforestry systems have been shown to reduce runoff and soil evaporation (Kinama et al., 2005) as well as increase soil fertility and create beneficial microclimatic effects (Mbow et al., 2014). In Central Ethiopia, ~200 km north of our study area, grain yield and aboveground biomass of wheat increased under the canopy of *F. albida* (Sida et al., 2018a). In this case, temperature regulation helped mitigate heat stress, promote higher water use efficiency, and increased soil mineral nitrogen, resulting in higher wheat yields aided by *F. albida* (Sida et al., 2018a).

With few exceptions, there has been little research on staple food crops demonstrating positive impacts of forests on yields. More commonly reported are instances where trees and forests negatively impact staple crop yields. For maize and soybean in Canada, net assimilation and photosynthetic radiation (PAR) (both highly correlated with crop yield), decreased within 2 m of intercropped trees (Reynolds et al., 2007). Similarly, wheat yield has generally been found to decrease with proximity to trees or forest in several other locations (Akbar et al., 1990; Puri and Bangarwa, 1992; Khan and Ehrenreich, 1994; Reynolds et al., 2007; Schmidt et al., 2019). However, it is ultimately external geographic, biological, or other environmental factors that determine whether trees and forest positively or negatively impact wheat. For example, extreme heat stress has been shown to drastically decrease wheat yield (Farooq et al., 2011; Lobell et al., 2012), leading to instances where wheat yield is higher in areas that are shaded and have lower temperatures (Sida et al., 2018b).

---

**Table 5** | $P$-values resulting from Tukey’s range test between the distances classes from linear mixed effect models for each vegetation index.

<table>
<thead>
<tr>
<th></th>
<th>NDVI</th>
<th>GCVI</th>
<th>MSAVI2</th>
</tr>
</thead>
<tbody>
<tr>
<td>p-value</td>
<td>p-value</td>
<td>p-value</td>
<td></td>
</tr>
<tr>
<td>10–30 vs. 31–50 m</td>
<td>0.1938</td>
<td>0.1039</td>
<td>0.0033</td>
</tr>
<tr>
<td>10–30 vs. 51–70 m</td>
<td>0.0166</td>
<td>0.0011</td>
<td>0.0162</td>
</tr>
<tr>
<td>10–30 vs. 71–90 m</td>
<td>0.0003</td>
<td>&lt;0.0001</td>
<td>0.0015</td>
</tr>
<tr>
<td>31–50 vs. 51–70 m</td>
<td>0.8244</td>
<td>0.3727</td>
<td>0.9952</td>
</tr>
<tr>
<td>31–50 vs. 71–90 m</td>
<td>0.2623</td>
<td>0.0472</td>
<td>0.9181</td>
</tr>
<tr>
<td>51–70 vs. 71–90 m</td>
<td>0.9309</td>
<td>0.9467</td>
<td>0.9949</td>
</tr>
<tr>
<td>Control vs. 10–30 m</td>
<td>0.0006</td>
<td>&lt;0.0001</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Control vs. 31–50 m</td>
<td>0.5511</td>
<td>0.0002</td>
<td>0.0053</td>
</tr>
<tr>
<td>Control vs. 51–70 m</td>
<td>0.9998</td>
<td>0.1901</td>
<td>0.0472</td>
</tr>
<tr>
<td>Control vs. 71–90 m</td>
<td>0.9284</td>
<td>0.4792</td>
<td>0.0705</td>
</tr>
</tbody>
</table>

All three VIs have 4,037 degrees of freedom. $P$-values $<0.05$ have been bolded.
Our results showing higher VI (a proxy for yield) in close proximity to the forest edge is especially interesting because the relationship between proximity to forest patches and wheat productivity can have important implications on desirable landscape configuration and food security in the region. Although the results of competition and facilitation between forests and agriculture may be mixed, our results are suggestive of some potential benefits which may counteract nutrient, light, and water competition from trees and forests. Although our study does not isolate the exact mechanisms behind this increase in productivity, it is likely that microclimatic effects (such as those demonstrated with *F. albida*) play a part (Laike, 1991; Sida et al., 2018a). If trees and forests can provide beneficial microclimatic buffering impacts, this could aid climate change mitigation strategies. Nonetheless, it is important to note that this benefit may change depending on geographic and biologically-dependent factors.

**Limitations and Additional Future Research**

Despite the high spatial resolution (5 and 10 m) of our imagery, our vegetation indices and resulting proximity analysis could have been unduly affected by mixed signals from nearby forests. Pixels at the margin of wheat fields (and thus closest to forest) could be impacted by spectral signatures of nearby trees and forests which would artificially inflate measured VI values. However, we took several measures to lessen this problem. Our VI values were derived from a coarser resolution Sentinel-2 (10 m) imagery while the classification of forests utilized finer resolution RapidEye (5 m) imagery. While this was originally done due to a lack of RapidEye availability during a specific time, when incorporated into the proximity analysis, this approach helped ensure that no wheat pixels located near the forest edge consisted of mixed forest pixels. Finally, we also tried to mitigate this problem within our sampling design by leaving a 10 m gap between the forest edge and first distance class (10–30 m) for our statistical comparisons.

Our study was also limited by a lack of image availability. High cloud cover during the rainy seasons and high commercial demand elsewhere (which translates to less archived imagery of the region) prevented the capture and archival of RapidEye images during several relevant time periods in the study area. A lack of imagery during pre-harvest season was one reason the land cover classification used to characterize the fields was from a different year and source (2015, RadpIdEye) than the image used to derive the VIs (2016, Sentinel-2). Additionally, the only cloud-free image available during the pre-harvest period in 2014 was for October 26th. Because many fields were already harvested before this date, only 11 out of the 23 fields were able to be analyzed to determine the VI-yield relationship.

The landscape configuration of our study area also created a limitation in the sampling design. In selecting the sample fields we made sure that a given treatment distance class was sufficiently far away from any surrounding non-treatment forests edges (at least 80 m). However, because average field size in the region was small (0.3 ha) and hedgerows were abundantly scattered throughout the landscape it was hard to find sample fields big enough where the distance classes were also 80 m away from edges of other fields. Therefore, our study did not account for non-forest field edges.

Two lines of future research are deemed particularly important to support and extend our research findings. First, we propose additional field work to conduct crop cuts to validate the edge effects on wheat productivity, along with exploring the potential mechanisms behind such edge effects (e.g., shade vs. nitrogen-fixing hedgerows). Second, we modeled different landscape configurations of forests and fields to explore regional landscape scenarios which maximize benefits to farmers (Yang et al., 2020).

**CONCLUSIONS**

Improving food security throughout the developing world remains one of humanity’s greatest challenges in the twenty first century. Of the many different facets affecting this complex problem, the first part of our research improves regional scale crop yield monitoring in smallholder farms. By analyzing VIs from accessible, low cost, high spatial resolution satellite imagery, we provide an alternative to conventional on-site field monitoring of crops on smallholder farms. With increasing availability of high and very high resolution data (e.g., Skysat, Pléiades, Worldview-3, etc.), it will be increasingly feasible to assess not only crop health but other variables affecting yields such as forests and forest edges.

The second half of this research investigated wheat productivity as it relates to nearby forest patches in Southern Ethiopia. VIs in wheat fields were significantly higher in field margins closer to forests. Although the exact mechanisms behind such VI increases is unknown, it is possible that microclimatic effects are a contributing factor. Both the grain filling and antithesis development stages of wheat are significantly affected by water availability and temperature (Fischer et al., 1998; Gooding et al., 2003). Heat stress, and potentially a lack of water availability via drought, are factors that could be mitigated by the effects of forests. With temperatures projected to increase significantly in the Central Rift Valley, adoption of agroforestry based practices or landscape approaches, including the retention of forest patches could mitigate some of the effects of climate change.

In addition, future climatic changes are likely to impact food security in many regions (Schmidhuber and Tubiello, 2007; Wheeler and von Braun, 2013). Higher temperatures and lower water availability could have severe impacts on crops such as wheat (Intergovernmental Panel on Climate Change, 2014). In fact, global wheat production is projected to decline significantly under current IPCC scenarios as a result of heat stress and drought (Parry et al., 2004; Lobell et al., 2011; Semenov and Shewry, 2011). Forests have been shown to create microclimates that alleviate extreme heat and help retain soil moisture (Camargo and Kapos, 1995; Gehlhausen et al., 2000; Chen et al., 2010). If forests can provide beneficial microclimatic
conditions, our research adds to the growing body of literature suggesting positive edge effects from forests despite some known potential negative effects.

**DATA AVAILABILITY STATEMENT**

The datasets generated for this study are available on request to the corresponding author.

**AUTHOR CONTRIBUTIONS**

FB conceived of the initial hypotheses and assisted with statistical analyses. KY wrote the manuscript and conducted the analyses. SG guided the methods and study design. J-YD-C was responsible for ground data collection and field-size mapping. All authors contributed to editing the manuscript and providing feedback that shaped the research, analysis, and interpretation.

**REFERENCES**


**FUNDING**

This study was conducted with funding from CRP WHEAT (www.wheat.org; C4032 Maize S12 Systems Intensification–CIMMYT subgrant, FB and SG) and NSERC–Discovery Grant (2014-05012, SG).

**ACKNOWLEDGMENTS**

This work was implemented by UBC and CIMMYT thanks to the support of CRP WHEAT (www.wheat.org). Any opinions, findings, conclusion, or recommendations expressed in this publication are those of the authors and do not necessarily reflect the view of CRP WHEAT. We would like to thank Dr. Nicholas Coops and Dr. Sean Smukler for their feedback and advice. We would also like to thank Kedir Gemedo for supporting data collection.


