Cover crop mapping in lower peninsula of Michigan State using Harmonized Landsat 8 and Sentinel-2 (HLS) data

Shao, Yiwen
4-30-2021

Abstract

Planting cover crops provides many benefits, including mitigating soil erosion, increasing soil fertility, and managing weeds, yet there is limited understanding of where and when cover crops have been planted. In this study, we use the Harmonized Landsat Sentinel-2 (HLS) surface reflectance data product to map winter land cover, including cover crop species, across three sites in the lower Michigan peninsula using random forest models. Our results showed a moderate overall accuracy (66%) across all three sites, with individual level accuracies varying by region and land cover type. Considering which bands and time periods were most important for classification, we found that vegetation indices developed using the red edge bands in the early part of the growing season were particularly important. This work suggests that readily-available satellite data can be used to accurately map cover crop species, though accuracies are lower than previous products that simply created binary classifications of cover crop presence or absence. Future work should examine whether classification accuracies can be improved with increased training data available for the most difficult to classify land cover classes, and by using Sentinel-2 data which may be able to provide red-edge data during the early part of the growing season due to an increased number of available clear-day scenes.

Introduction

The world's population is expected to increase over the coming decades [1], and studies estimate that food production will have to increase by 70-100% in order to meet growing demand [2]. While yields have historically increased through agricultural intensification, the increased use of pesticides, inorganic fertilizers, and freshwater for irrigation has led to significant environmental externalities, including soil erosion, groundwater depletion, and eutrophication [3,4]. Such negative environmental externalities are particularly acute in the Midwestern United States, where excess nitrogen runoff has led to hypoxic conditions in the Gulf of Mexico and harmful algal blooms in the Great Lakes [5,6]. Such undesirable consequences of conventional agriculture have prompted the adoption of more sustainable agricultural practices, including crop rotations, reduced tillage, organic fertilizers, and cover crops [7]. Cover crops in particular can offer a wide range of benefits. Cover crops have been shown to mitigate erosion, reduce surface runoff, improve soil infiltration, retain soil moisture, and suppress weeds [8]. Despite such benefits, the adoption of cover crops among U.S. farmers has historically been low (< 5% of national cropland in 2012) [9], though there is evidence that the area under cover crops is increasing through time [10]. Yet to date it has been challenging to track where and when cover crops have been adopted, and whether farmers continue to use cover crops through time as it is challenging to collect spatially-explicit data annually through typical methods such as censuses.

Remote sensing offers a viable way to map cover crop adoption and use at large spatial and temporal scales at low cost [11]. Previous studies have used Landsat and Sentinel-2 data to map cover crops in multiple regions across the globe, including the eastern and Midwest United States, Japan, and Spain [12,13,14]. While existing studies have mapped cover crop extent with high accuracies, they have focused on conducting binary classifications to determine the presence of any cover crop species. To date, studies have not attempted to map specific cover

crop species at the landscape scale using satellite data. Yet, identifying which specific cover crop species have been planted and where is important given that different cover crop species can have different impacts on soil health, nitrogen runoff, and crop yields [15,16]. Furthermore, understanding where specific cover crop species have been planted can better reveal potential mechanisms driving cover crop adoption as farmers may prefer to plant different species for different reasons [17,18].

Previous studies that have mapped cover crops at scale have largely used Landsat and Sentinel-2 imagery. Sentinel-2 imagery is at a higher spatial resolution (10 m) and also contains additional spectral bands, such as red edge bands, compared to Landsat satellite data (30 m resolution). Red-edge bands in particular have been shown to be important when mapping crop type as they capture information in a part of the spectrum that can be used to identify vegetation due to its unique spectral signature in this region [19]. Previous studies have shown that using Sentinel-2 red-edge bands could help distinguishing between different crop types, suggesting they may be helpful in distinguishing between different cover crop types [20,21,22]. The temporal resolution of Sentinel-2 (5 days), however, is lower than that of Landsat (2-3 days). Previous studies have shown that high temporal information can be important for identifying crop types as it can help distinguish different phenologies for different crops [21]. The Harmonized Landsat Sentinel (HLS) data product combines Landsat and Sentinel-2 satellite data by aligning both data products on the same tiling system and aggregating Sentinel-2 data to the spatial resolution of Landsat. This data product provides global measurements at 30 m spatial resolution every 2-3 days. HLS may offer advantages over using individual sensors as it combines most of the key benefits of each sensor into one data product.

In this study, we used the HLS dataset to classify winter cover, including cover crops, in lower Michigan. We specifically ask the following research questions:

1) How effectively can HLS map winter cover, including cover crops, across multiple regions in Michigan?
2) Can we develop a generalizable algorithm that can map winter cover accurately across multiple sites with varying climate, farm management practices, and soil types?
3) Which bands, indices, and time periods are most important for classifying winter cover?

Our study provides important insights into the ability of readily-available satellite imagery to map cover crops across Michigan, and the Midwestern United States more broadly. Mapping cover crop adoption across this region is of importance given that it is one of the world's hotspots for nitrogen pollution and understanding the extent and consistency of cover crop adoption can help better understand how effective cover crops may be in reducing leaching and eutrophication of associated waterways.

2. Methods

2.1 Study area

The study area is divided into three regions in the lower peninsula of Michigan: the southwest region (SW) spreads over three counties (Van Buren, St. Joseph, and a portion of Berrien), the southeast region (SE) contains two counties (Lenawee and Monroe), and the thumb region (TB) comprises five counties (portions of Huron, Tuscola, Sanilac, Lapeer, and St. Clair;

Figure 1). We selected these three regiond because they contain counties with relatively large land areas planted unfrt cover crops according to the 2012 and 2017 USDA census of agriculture (e.g., from 5-27% of agricultural land in 2017), and span contrasting climate conditions, soil types, and management systems. Prior to initiating data collection, we met with extension agents from MSU and district conservationists in each region to finalize the choice of study sites. The two clusters in southern Michigan have high-input, row crop production systems that are significant sources of the N and P losses that cause eutrophication of the Great Lakes. The southwestern region has sandier soil types, while the southeastern region has heavier, clay soils. The Thumb region of Michigan has a growing presence of large-scale grain farms, but also has dairy production and the largest cluster of organic grain farmers in the state. Given that cover crops are relatively rare on Midwestern agricultural landscapes, we selected areas where we would be more likely to find multiple cover crop types, but within regions where simplified production systems are common.

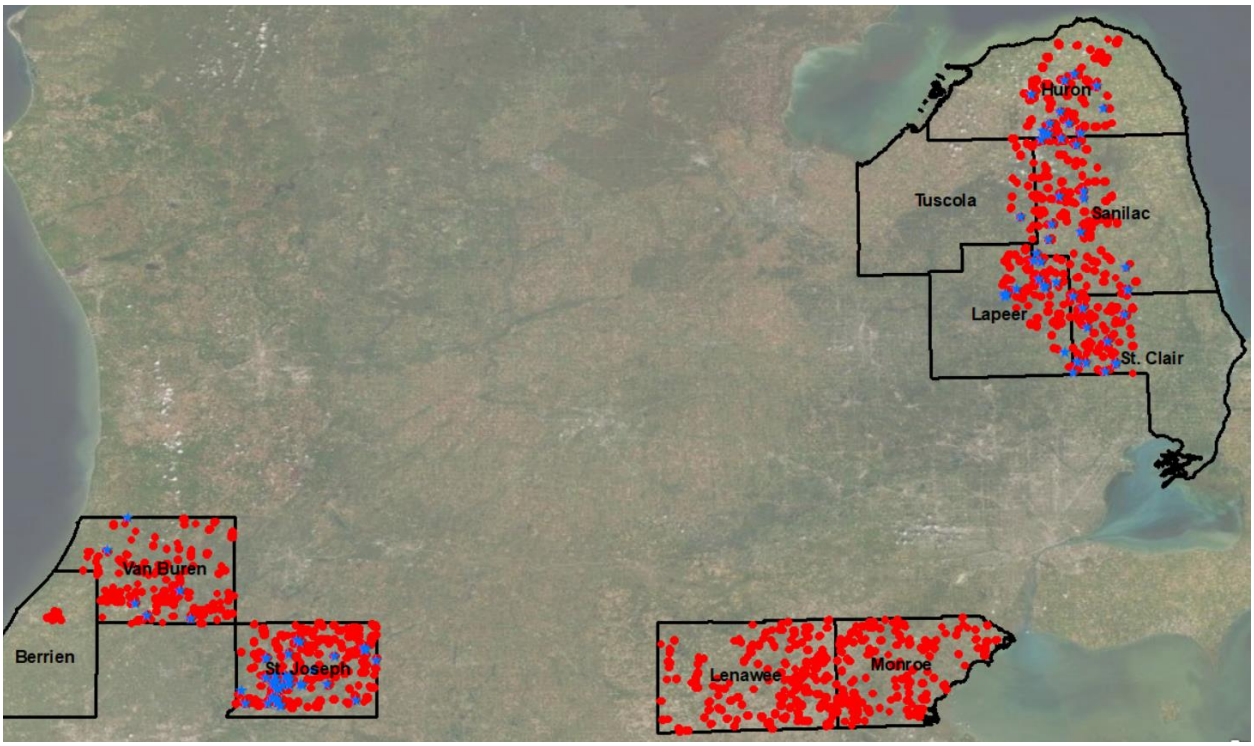

Figure 1: The three study regions, Southeast (SE), Southwest (SW), and the Thumb (TB). The red symbols denote the fields visited in April-June 2019 and blue symbols are the fields re-visited in July to distinguish between winter wheat and cereal rye cover crop fields.

2.2 Field data

We conducted field surveys in three stages across the three regions during April to July 2019. During the first stage of data collection (April 24 – May 3), we visited known cover crop locations using information provided by extension agents, researchers, and farmers. Extension contacts in each region shared information about our project through e-lists and newsletters, and farmers submitted field addresses or GPS points of fields with annual cover crops. For all three stages, data were collected using either handheld GPS units (Garmin eTrex 20), or the ArcCollector app (https://www.esri.com/en-us/arcgis/products/collector-for-arcgis/overview).

Each field was classified by visual interpretation of the existing landcover type, with crops identified and assigned confidence levels to indicate certainty of identification (scale of 1 to 5, 5 = highly certain). We listed all crops that were visible in mixed planted fields. We took photos of any fields classified as low to moderate certainty for reference.

During the second stage of data collection (May 3 – June 2), we revisited all three regions to collect GPS points systematically for all types of winter plant cover. Specifically, field teams visited 320 pre-selected points per region, which were spatially distributed throughout each region. These points were selected using a stratified random sampling approach. First, we obtained Sentinel-2 imagery for the start of the cover crop growing season (from April 1-28, 2019) prior to our field survey, and calculated maximum Normalized Difference Vegetation Index (NDVI) for each pixel. Next, we used the 2018 cultivated layer from the USDA, National Agricultural Statistics Service (NASS) Cropland Data Layer (CDL) (USDA NASS, 2017) to mask out non-agricultural areas. We then examined histograms of NDVI for the unmasked pixels across our three study regions, and binned NDVI values into four relatively equal categories (< 0.2, 0.2 - 0.4, 0.4 - 0.6 and > 0.6). We next used a road layer and selected 80 points per NDVI class that were within 50 meters of a road, resulting in 320 points per study region. Point sampling was conducted using ArcGIS software.

In the third stage of data collection during the first week of July (July 1 – July 7), we revisited and confirmed field classifications for cereal rye and winter wheat, given that these two crops were similar in appearance during the first two stages of data collection. Figure 1 depicts the fields surveyed during the first two stages as red symbols and the fields verified in the third stage as blue symbols. We overlaid the GPS data on high-resolution Google Earth imagery and manually digitized field boundaries using visual interpretation of field edges. We excluded field polygons where the confidence level of crop identification was 3 and below. The total number of fields that remained were 328 in SE, 414 in SW, and 350 in TB (Figure 2). We examined the NDVI phenologies for each cover type (Figure 3), and found that the NDVI phenologies were very similar for bare (B) and tilled (T) fields. Therefore, we aggregated these two classes together into one class bare/tilled (BT) for all analyses.

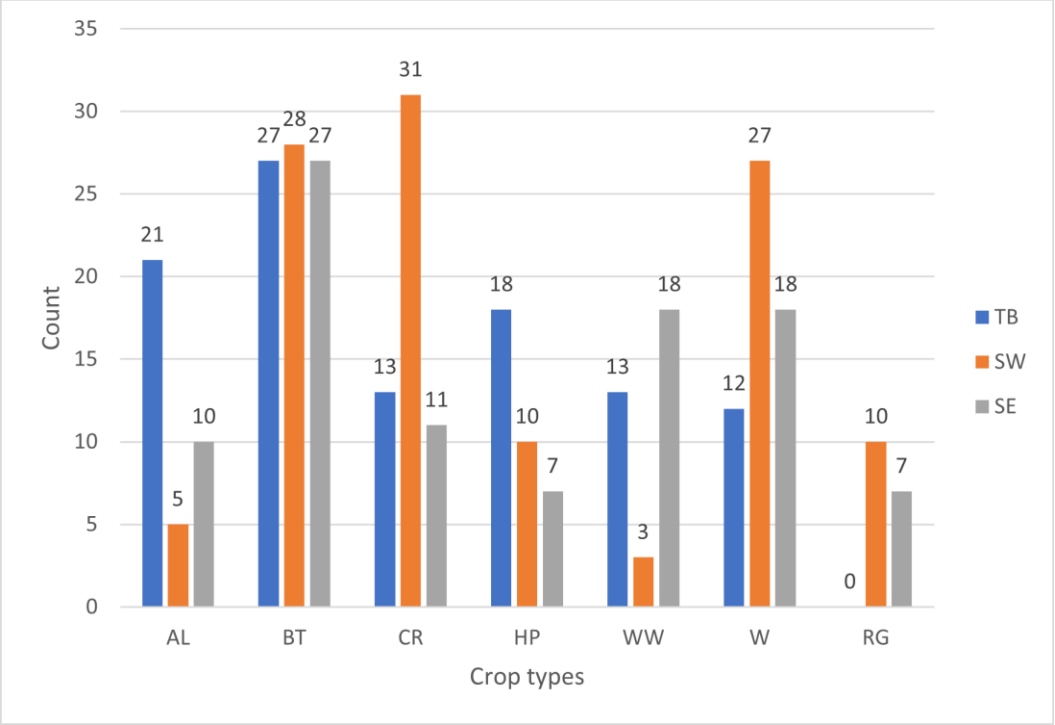

Figure 2: Bar plots showing the counts of cover crops in sampled fields of the three regions. AL = alfalfa, BT= bare/tilled, CR = cereal rye, HP = hay forage/pasture, WW=winter wheat, W = weeds, RG = ryegrass.

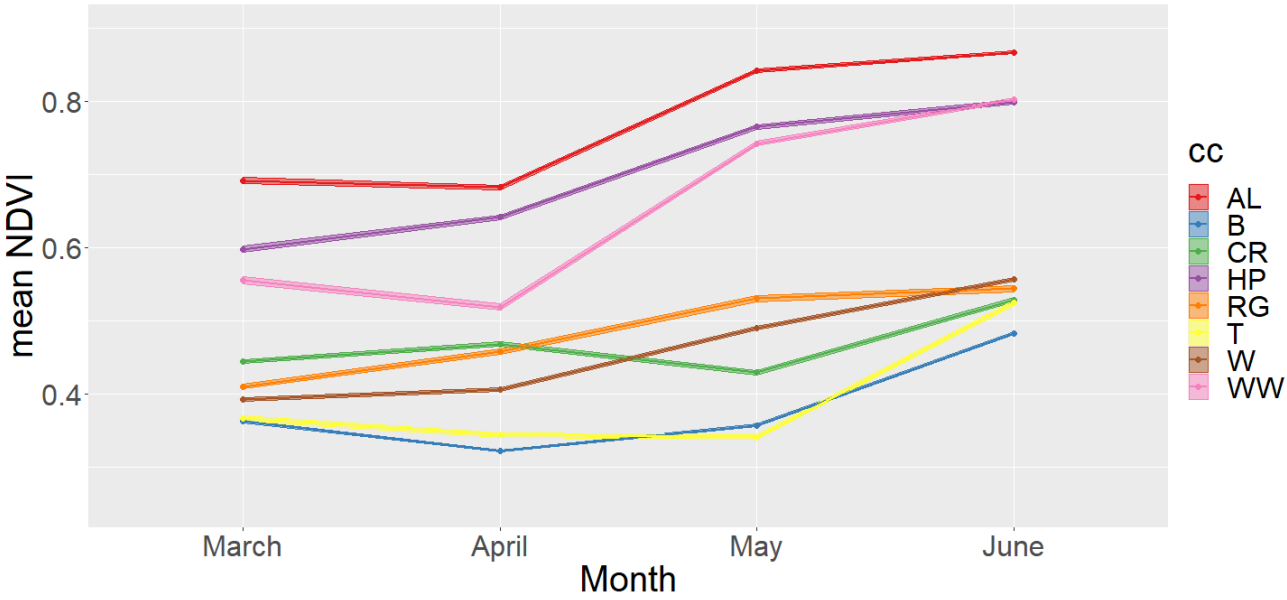

Figure 3: Mean NDVI phenology with 95% confidence intervals across all available polygons of different land cover types. AL = alfalfa, B= bare, T= tilled, CR = cereal rye, HP = hay forage/pasture, WW=winter wheat, W = weeds, RG = ryegrass. Month 3 represents March, and month 6 represents June.

2.3 Satellite Data and Preprocessing

We processed the Harmonized Landsat Sentinel-2 product (HLS) from March 1st to June 30th, 2019 to encompass the main growing period for cover crops in our study region. The HLS data were downloaded from the HLS website (https://hls.gsfc.nasa.gov/) using the available batch script. We specifically downloaded the S30 data product (Sentinel-2 Multispectral Instrument (MSI) surface reflectance resampled to 30 m in the Sentinel-2 tiling system and adjusted to Landsat 8 spectral reflectances) and the L30 data product (Landsat OLI harmonized surface reflectance and TOA brightness temperature resampled to 30 m in the Sentinel-2 tiling system). The spectral bands for S30 and L30 are the same as those for the Sentinel-2 MSI and Landsat-8 OLI data products, respectively [23,24]. The cloud mask we used for the L30 product is based on the Quality Assessment (QA) layer. For the S30 data product, we did not use the QA layer as previous work and examination of our data show that it is inaccurate, often labeling non-cloudy pixels as cloud covered [23]. Instead, we adapted a Sentinel-2 cloud mask algorithm developed in Google Earth Engine by Ian Housman (https://groups.google.com/g/google-earth-engine-developers) and applied it in R project software. We then created monthly mosaics using L30 and S30 data for all bands that were common between the two datasets, and only S30 data for those bands that were specific to Sentinel-2. To create monthly mosaics, we selected the pixel that had the highest NDVI value across all available values within the month, and extracted all available bands for that image. We did not, however, extract data for the Coastal Aerosol and Cirrus bands as these do not measure surface reflectance characteristics. We found that there were many missing pixels in the S30 monthly mosaics for March and May. Therefore, we did not include S30 only bands for these two months in our final dataset. We calculated twelve spectral indices that have been shown to be important for monitoring agricultural crop characteristics in the previous literature (Table 1). All cloud masking, mosaicking, and index creation were performed using R Project software [25] and the ‘sf’ [26], ‘raster’[27], and ‘caret’[28] packages.

Table 1. Spectral bands and indices included in our mosaics

	Products	
Spectral Bands	HLS L30	HLS S30
Blue	Band02	B02
Green	Band03	B03
Red	Band04	B04
Red-Edge 1	-	B05
Red-Edge 2	-	B06
Red-Edge 3	-	B07
NIR Narrow	Band05	B08A
SWIR 1	Band06	B11
SWIR 2	Band07	B12
Spectral Indices	Equation	
NDVI	(NIR - R) / (NIR + R)	
GBNDVI	(NIR - (G + B)) / (NIR + (G + B))	
GRNDVI	(NIR - (G + R)) / (NIR + (G + R))	
NPCI	(R - B) / (R + B)	

NDTI	(SWIR1 - SWIR2) / (SWIR1 + SWIR2)	
NDSVI	(NIR - G) / (NIR + G)	
GCVI	(NIR / GREEN) - 1	
SIWSI1	(NIR - SWIR1) / (NIR + SWIR1)	
SIWSI2	(NIR - SWIR2) / (NIR + SWIR2)	
NDI		(RE1 - R) / (RE1 + R)
PSRI		(R - G) / RE2
CIRE		RE3 / RE1 - 1

2.4 Sampling and Classification

We used random forest (RF), an ensemble tree-based classifier, to classify winter cover in this study [29]. We created separate models for each of the three regions (TB, SE, and SW), and one combined model for all regions together. We separated our ground truth polygons into 70% used for training and 30% used for validation across all models. To ensure that differences in field size did not affect our results, we selected 20 pixels at random from each polygon to use in our analyses, and these points were used consistently across all models. Before training the model, we used the findCorrelation function in the caret package [28] in R to remove highly-correlated variables ($r > 0.9$). We then used GEE to extract band values for all sensors and for all pixels, and exported these data for further analyses in R.

The random forest models were run using consistent parameters across all models, including using 10-fold cross validation for selecting training data, setting the 'mtry' parameter equal to the square root of the number of predictor variables included in the model, and setting 'ntree' equal to the default value of 500 [30, 31]. We ran random forest models using the caret [28] package in R project software [22]. We then used each model to predict winter land cover for all pixels in the validation dataset, and conducted validation at both the pixel and polygon scale. We classified each polygon using the majority class across all 20 selected pixels for each polygon. We used common metrics for validation, including overall accuracy, producer's accuracy, user's accuracy, and F1 scores [31, 32, 33]. In addition, we used the caret [28] package in R project software [22] to assess variable importance for each of the random forest models.

3 **Results**

3.1 Model accuracy

For the full study region (MI), the overall pixel-level accuracy was 62.82% and the overall polygon-level accuracy was 65.93% (Table 2, Table S1). Among the three individual region models, the Thumb (TB) (Table S2) and Southwest (SW) (Table S3) regions had similar accuracies, with pixel-based accuracies of 62.81% and 65.22%, respectively, and polygon-based accuracies of 67.31% and 68.29%, respectively (Table 2). The Southeast (SE) (Table S4) region had the lowest pixel and polygon-level accuracy of only 55.70% and 58.16%, respectively (Table 2). Accuracies for all regions improved when conducting validation at the polygon level compared to the pixel level.

Table 2. Overall accuracies

Accuracies	Regions			
	MI	TB	SE	SW
Pixel Accuracy	62.82%	62.81%	55.70%	65.22%
Polygon Accuracy	65.93%	67.31%	58.16%	68.29%

Table 3 shows the F1 scores for each land use type for each random forest model. In the generalized all site model, alfalfa (AL), bare/tilled (BT), cereal rye (CR), and winter wheat (WW) had the highest F1 scores. These patterns, however, were not consistent across individual site models. In particular, CR had high classification accuracy in the SW region, but low classification accuracies in the TB and SE regions. AL also had only modest accuracy in the SW region. Finally, WW had only modest accuracies in the SE and SW regions. Considering F1 scores for the generalized all site model, rye grass (RG) and weeds (W) had the lowest classification accuracies, and these classes showed low or modest accuracy in the individual site models.

Table 3. F1 scores for winter cover crops

Data set	Region	AL	BT	CR	HP	RG	W	WW
HLS	MI	0.744	0.740	0.751	0.571	0.300	0.523	0.829
	TB	0.750	0.746	0.364	0.718		0.333	0.870
	SE	0.800	0.742	0.154	0.000	0.600	0.419	0.647
	SW	0.500	0.720	0.824	0.333	0.625	0.627	0.667

3.2 Important predictors

Considering variable importance, there were several trends across models (Table 4, Figures S1-S4). Vegetation indices appeared as the top five most important variables across all models, with only two individual bands (NIR and green) appearing as the most important variables for some models. In addition, most important variables were from earlier in the growing season, particularly from the months of April and May. Images from June appeared to be more important for the SE region (Figure S4), as two of the five most important variables were from June. We found that six of twenty most important variables across all models include vegetation indices that use the red edge bands, highlighting the importance of these spectral bands for classification (Table 4).

Table 4. Top 5 Important Variables for Classification Models

Region	Top Important Variables (Month)				
	1	2	3	4	5
MI	NDTI (May)	GCVI (May)	NIR (May)	CIRE (April)	PSRI (April)
TB	GCVI (May)	NIR (May)	GCVI (June)	PSRI (April)	SWIR1 (May)
SW	NDTI (May)	PSRI (April)	NIR (May)	CIRE (April)	NDTI (April)
SE	GCVI (May)	Green (June)	PSRI (April)	GBNDVI (May)	GCVI (June)

4 Discussion and Conclusion

We used HLS satellite data to map winter cover, including cover crop species, across lower Michigan. We find that our random forest models produce moderate overall accuracies (66% across all three sites) across Michigan and in each site except for the Southeast site, where accuracies were 8% lower (58%). The classification accuracy varied across landcover classes, with the highest overall accuracies achieved for winter wheat, alfalfa, and bare/tilled fields. There was substantial variation in performance across land cover classes across sites, with some classes such as cereal rye performing well in some sites (82% in the Southwest site) and very poorly in others (15% in the Southeast site). Overall, we find that HLS data are able to map winter cover with moderate accuracies, and individual cover crop species with low to moderate accuracies depending on species and region.

There was significant variation in overall accuracy across sites, with the Southwest site having the highest accuracy (68%), followed by the Thumb site (67%), and then followed by the Southeast site (58%; Table 3). There are several reasons that may explain this difference in accuracy. The Southeast site had relatively fewer training polygons for most land cover classes, other than winter wheat, compared to the other two sites (Figure 2). In addition, land management in the Southeast is more heterogeneous than in the other two regions, which could have made it difficult for the random forest models to classify each class accurately. The Thumb site may have had higher classification accuracies because the site did not contain any rye grass (Figure 2), which was a difficult land cover to classify (Table 4). Finally, the Southwest site may have had higher overall accuracies because classification of difficult land cover classes, such as cereal rye and weeds, had moderate to high classification accuracies in this site. This may be because we were able to get increased training data for these two difficult landcover classes in the Southwest region (Figure 2) [35], likely due to higher prevalence across the landscape. Future work should attempt to disentangle the role that the number of training data may play in variation in classification accuracies by collecting the exact same number of training data for all land cover classes in each region [35,37].

Considering which bands and time periods were most important for classification, we found that vegetation indices developed using the red edge bands were particularly important (Table 4). Six of the top 20 most important variables across all four models contained red edge bands. This is not surprising given that previous studies have shown that the red edge is particularly useful in classifying different vegetation types [38,39]. Considering time period, our results suggest that images from earlier in the growing season, particularly April and May, were most important for classifying winter cover. This makes sense given that the phenology plots that we created for each of the landcover types considered in our study showed the greatest differences during April and May (Figure 4). Unfortunately we were unable to include red edge bands for March and May due to high amounts of cloud cover in the S30 product, and it is possible that accuracies could improve if red edge bands for these months were available. Based on our examination of S30 data in comparison to Sentinel-2 Level-2A data, we found that there were many clear day scenes that were available in the Sentinel-2 product that were not available using the S30 data product. It is possible that if we were to use Sentinel-2 Level-2A data directly, we may have had enough clear scenes available to create a viable mosaic of red edge bands. Future work should examine how well Sentinel-2 data can be used to classify winter cover in this region.

In conclusion, we found that HLS data were able to map winter land cover, including cover crops, with moderate accuracy across the three sites considered across lower Michigan. This

work suggests that readily-available satellite data can be used to accurately map cover crop species, though accuracies are lower than previous products that simply created binary classifications of cover crop presence or absence. Future work should examine whether classification accuracies can be improved with increased training data available for the most difficult to classify land cover classes, and by using Sentinel-2 data which may be able to provide red-edge data during the early part of the growing season.

References

1. Coole, D. (2018). Should We Control World Population?. United Kingdom: Wiley.
2. McKenzie, F. C., & Williams, J. (2015). Sustainable food production: constraints, challenges and choices by 2050. Food Security, 7(2), 221-233.
3. Kessler, J. J., & Moolhuijzen, M. (1994). Low external input sustainable agriculture: expectations and realities. NJAS wageningen journal of life sciences, 42(3), 181-194.
4. Verma, A. K. (2018). Unsustainable agriculture, environmental ethics and ecological balance. HortFlora Research Spectrum, 7(3), 239-241.
5. Rabalais, N. N., & Turner, R. E. (2019). Gulf of Mexico hypoxia: Past, present, and future. Limnology and Oceanography Bulletin, 28(4), 117-124.
6. Kerr, J. M., DePinto, J. V., McGrath, D., Sowa, S. P., & Swinton, S. M. (2016). Sustainable management of Great Lakes watersheds dominated by agricultural land use. Journal of Great Lakes Research, 42(6), 1252-1259.
7. Lu, Y. C., Watkins, K. B., TEASDALE, J. R., & ABDUL-BAKI, A. A. (2000). Cover crops in sustainable food production. Food Reviews International, 16(2), 121-157.
8. Teasdale, J. R. (1996). Contribution of cover crops to weed management in sustainable agricultural systems. Journal of production agriculture, 9(4), 475-479.)
9. U.S. Department of Agriculture. (2012) Census of Agriculture 2012, Volume 1
10. CTIC. (2020). ANNUAL REPORT 2019-2020 national cover crop survey. Conservation Technology Information Center, Sustainable Agriculture Research and Education Program, and the American Seed Trade Association. West Lafayette, IN.
11. Hively, W. D., Duiker, S., McCarty, G., & Prabhakara, K. (2015). Remote sensing to monitor cover crop adoption in southeastern Pennsylvania. Journal of Soil and Water Conservation, 70(6), 340-352.
12. Seifert, C. A., Azzari, G., & Lobell, D. B. (2018). Satellite detection of cover crops and their effects on crop yield in the Midwestern United States. Environmental Research Letters, 13(6), 064033.
13. Sonobe, R., Yamaya, Y., Tani, H., Wang, X., Kobayashi, N., & Mochizuki, K. I. (2017). Mapping crop cover using multi-temporal Landsat 8 OLI imagery. International Journal of Remote Sensing, 38(15), 4348-4361.
14. Cruz-Ramírez, M., Hervás-Martínez, C., Jurado-Expósito, M., & López-Granados, F. (2012). A multi-objective neural network based method for cover crop identification from remote sensed data. Expert Systems with Applications, 39(11), 10038-10048.
15. Frasier, I., Quiroga, A., & Noellemeyer, E. (2016). Effect of different cover crops on C and N cycling in sorghum NT systems. Science of The Total Environment, 562, 628-639.
16. Demir, Z., Tursun, N., & Işık, D. (2019). Effects of different cover crops on soil quality parameters and yield in an apricot orchard. International Journal of Agriculture and Biology, 21(2), 399-408.
17. Wayman, S., Kucek, L. K., Mirsky, S. B., Ackroyd, V., Cordeau, S., & Ryan, M. R. (2017). Organic and conventional farmers differ in their perspectives on cover crop use and breeding. Renewable agriculture and food systems, 32(4), 376.
18. Roesch-McNally, G. E., Basche, A. D., Arbuckle, J. G., Tyndall, J. C., Miguez, F. E., Bowman, T., & Clay, R. (2018). The trouble with cover crops: Farmers' experiences with overcoming barriers to adoption. Renewable Agriculture and Food Systems, 33(4), 322-333.

19. Filella, I., & Penuelas, J. (1994). The red edge position and shape as indicators of plant chlorophyll content, biomass and hydric status. International journal of remote sensing, 15(7), 1459-1470.
20. Forkuor, G., Dimobe, K., Serme, I., & Tondoh, J. E. (2018). Landsat-8 vs. Sentinel-2: examining the added value of sentinel-2's red-edge bands to land-use and land-cover mapping in Burkina Faso. GIScience & remote sensing, 55(3), 331-354.
21. Ustuner, M., Sanli, F. B., Abdikan, S. A. Y. G. I. N., Esetlili, M. T., & Kurucu, Y. U. S. U. F. (2014). Crop type classification using vegetation indices of rapideye imagery. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 40(7), 195.
22. Radoux, J., Chomé, G., Jacques, D. C., Waldner, F., Bellemans, N., Matton, N., ... & Defourny, P. (2016). Sentinel-2's potential for sub-pixel landscape feature detection. Remote Sensing, 8(6), 488.
23. Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J.-C., Skakun, S. V., & Justice, C. (2018). The Harmonized Landsat and Sentinel-2 surface reflectance data set. Remote Sensing of Environment, 219, 145-161.
24. Landsat-8 image courtesy of the U.S. Geological Survey
25. R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
26. Pebesma, E. (2018). Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal 10 (1), 439-446, https://doi.org/10.32614/RJ-2018-009
27. Robert J. Hijmans (2019). raster: Geographic Data Analysis and Modeling. R package version 3.0-7. https://CRAN.R-project.org/package=raster
28. Max Kuhn (2020). caret: Classification and Regression Training. R package version 6.0-86. https://CRAN.R-project.org/package=caret
29. Breiman, L., & Cutler, A. (2004). RFtools—for predicting and understanding data. In Interface'04 Workshop.
30. Belgiu, M. and Dragut, L. (2016) Random Forest in Remote Sensing: A Review of Applications and Future Directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24-31. https://doi.org/10.1016/j.isprsjprs.2016.01.011
31. Lebourgeois, V., Dupuy, S., Vintrou, É., Ameline, M., Butler, S., & Bégué, A. (2017). A combined random forest and OBIA classification scheme for mapping smallholder agriculture at different nomenclature levels using multisource data (simulated Sentinel-2 time series, VHRS and DEM). Remote Sensing, 9(3), 259.
32. Bargiel, D. (2017). A new method for crop classification combining time series of radar images and crop phenology information. Remote sensing of environment, 198, 369-383.
33. Lantz, B. (2013). Machine learning with R. Packt publishing ltd.
34. Lawrence, R. L., Wood, S. D., & Sheley, R. L. (2006). Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). Remote Sensing of Environment, 100(3), 356-362.
35. Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. International Journal of Remote Sensing, 39(9), 2784-2817.
36. Fassnacht, F. E., Hartig, F., Latifi, H., Berger, C., Hernández, J., Corvalán, P., & Koch, B. (2014). Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. Remote Sensing of Environment, 154,

102-114.
37. He, H., & Garcia, E. A. (2009). Learning from imbalanced data. IEEE Transactions on knowledge and data engineering, 21(9), 1263-1284.
38. Schuster, C., Förster, M., & Kleinschmit, B. (2012). Testing the red edge channel for improving land-use classifications based on high-resolution multi-spectral satellite data. International Journal of Remote Sensing, 33(17), 5583-5599.
39. Clevers, J. G., & Gitelson, A. A. (2013). Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on Sentinel-2 and-3. International Journal of Applied Earth Observation and Geoinformation, 23, 344-351.
40. Chakhar, A., Ortega-Terol, D., Hernández-López, D., Ballesteros, R., Ortega, J. F., & Moreno, M. A. (2020). Assessing the accuracy of multiple classification algorithms for crop classification using Landsat-8 and Sentinel-2 data. Remote Sensing, 12(11), 1735.

Supplementary

Table. S1 HLS Confusion matrix for MI region (polygon)

Polygon		Reference							Total	User Accuracy
		AL	BT	CR	HP	RG	W	WW		
Prediction	AL	29	0	1	9	0	2	1	42	0.690
	BT	1	77	17	0	4	24	0	123	0.626
	CR	0	3	24	1	1	0	1	30	0.800
	HP	4	1	1	18	2	2	0	28	0.643
	RG	0	0	0	0	3	1	0	4	0.750
	W	1	4	6	5	6	29	2	53	0.547
	WW	1	0	5	2	0	0	29	37	0.784
	Total	36	85	54	35	16	58	33		
	Producer Accuracy	0.806	0.906	0.444	0.514	0.188	0.500	0.879		0.659

Table. S2 HLS Confusion matrix for TB region (polygon)

Polygon		Reference						Total	User Accuracy
		AL	BT	CR	HP	W	WW		
Prediction	AL	15	0	0	3	1	0	19	0.790
	BT	0	25	8	0	6	1	40	0.615
	CR	0	3	4	1	0	1	9	0.444
	HP	5	0	0	14	2	0	21	0.667
	W	1	0	0	0	3	1	5	0.600
	WW	0	0	1	0	0	10	11	0.909
	Total	21	27	13	18	13	12		
	Producer Accuracy	0.714	0.889	0.308	0.778	0.250	0.769		0.673

Table. S3 HLS Confusion matrix for SW region (polygon)

Polygon		Reference							Total	User Accuracy
		AL	BT	CR	HP	RG	W	WW		
Prediction	AL	3	0	0	3	0	0	0	6	0.500
	BT	1	27	3	0	2	10	0	43	0.628
	CR	1	3	28	1	1	2	0	36	0.778
	HP	1	0	1	3	0	1	1	7	0.429
	RG	0	0	0	1	5	0	0	6	0.833
	W	0	2	0	2	2	16	0	22	0.727
	WW	0	0	0	1	0	0	2	3	0.667
	Total	6	32	32	11	10	29	3		
	Producer Accuracy	0.500	0.844	0.875	0.273	0.500	0.552	0.667		0.683

Table. S4 HLS Confusion matrix for SE region (polygon)

Polygon		Reference							Total	User Accuracy
		AL	BT	CR	HP	RG	W	WW		
Prediction	AL	10	0	0	5	0	0	0	15	0.667
	BT	0	23	2	0	1	9	0	35	0.657
	CR	0	1	1	0	0	0	0	2	0.500
	HP	0	0	1	0	0	0	1	2	0.000
	RG	0	0	0	0	3	0	0	3	1.000
	W	0	3	3	2	2	9	6	25	0.360
	WW	0	0	4	0	1	0	11	16	0.688
	Total	10	27	11	7	7	18	18		
	Producer Accuracy	1.000	0.852	0.091	0.000	0.429	0.500	0.611		0.582

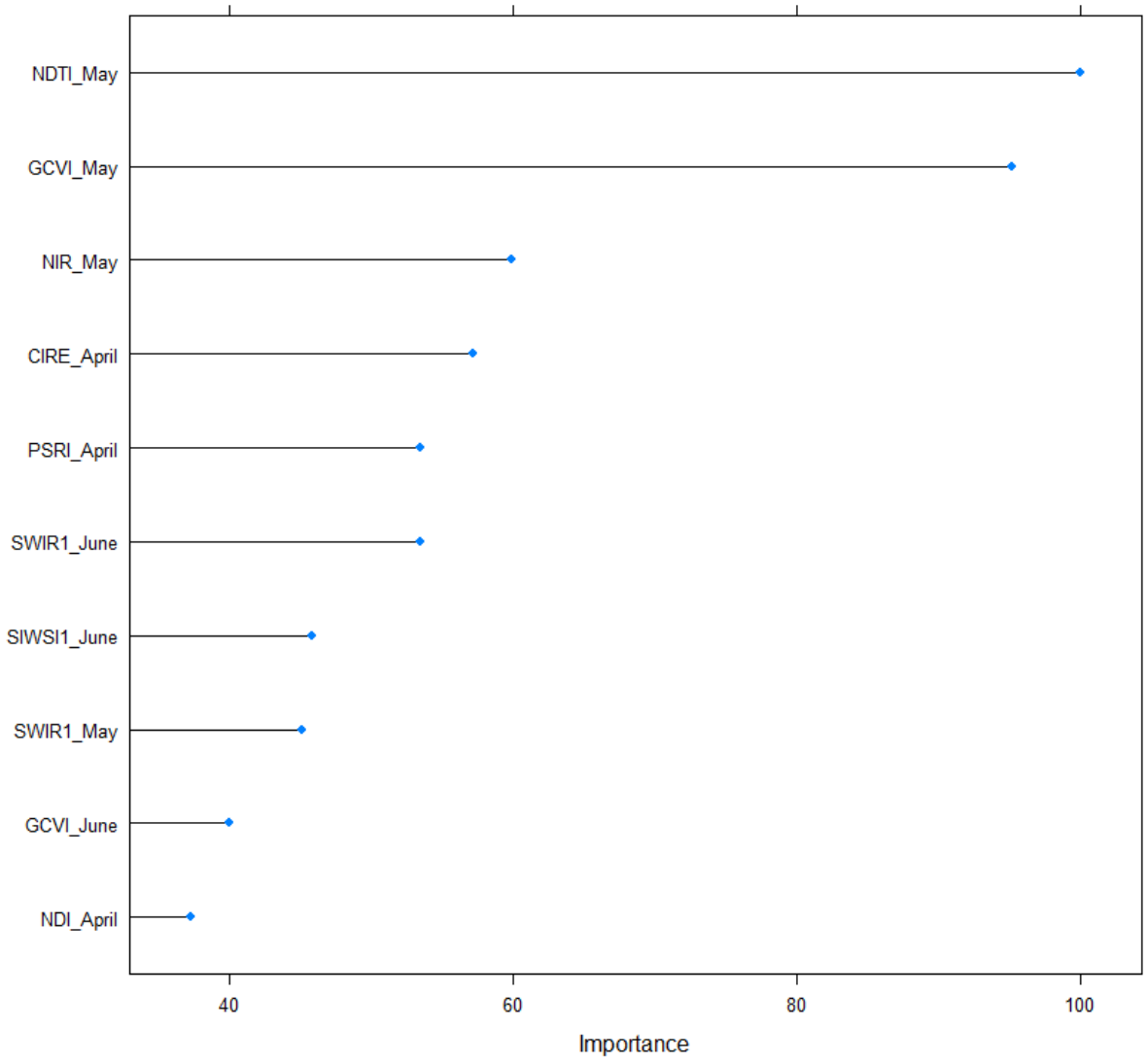

Figure S1. Variable Importance for MI region (HLS)

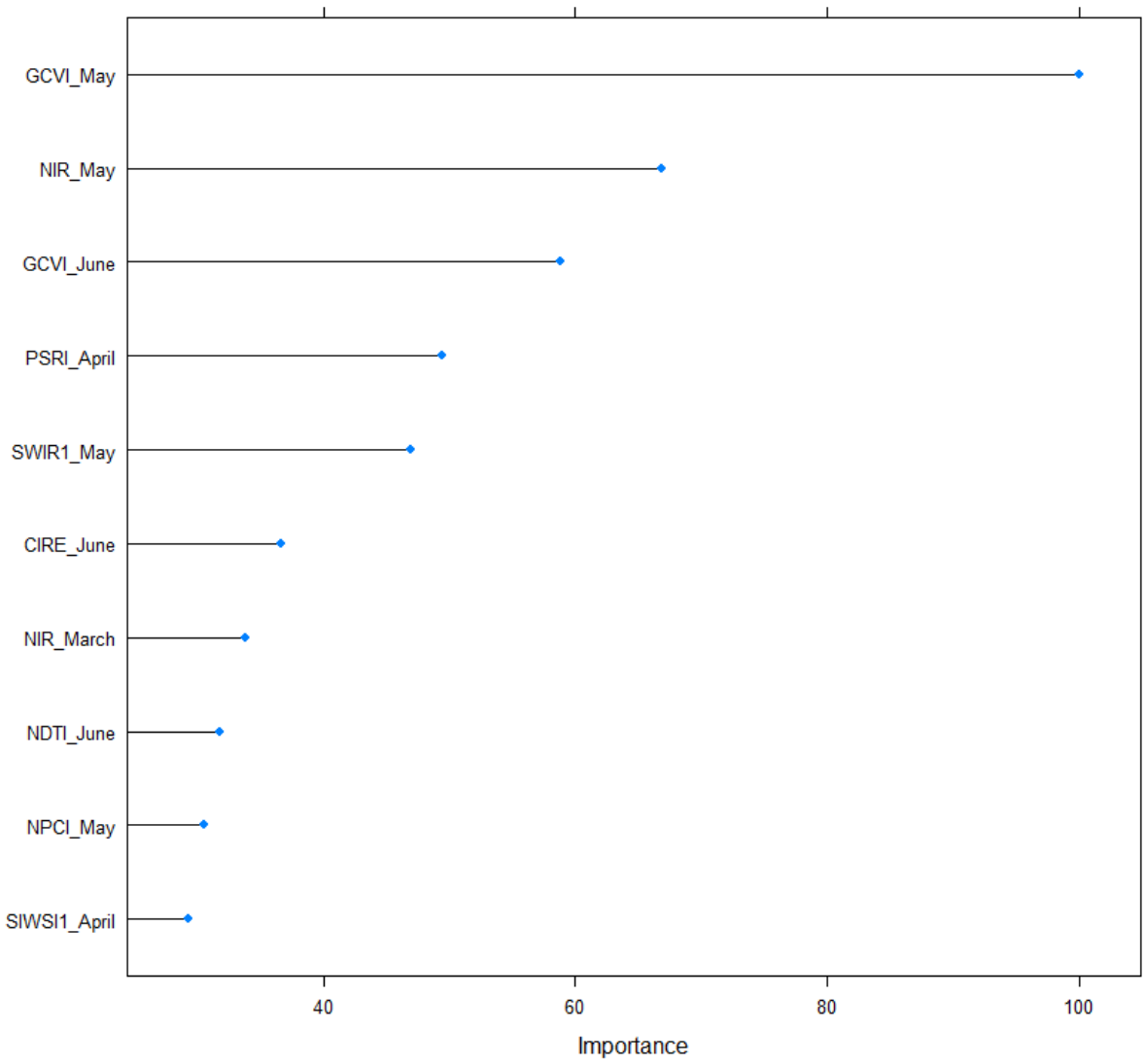

Figure S2. Variable Importance for TB region (HLS)

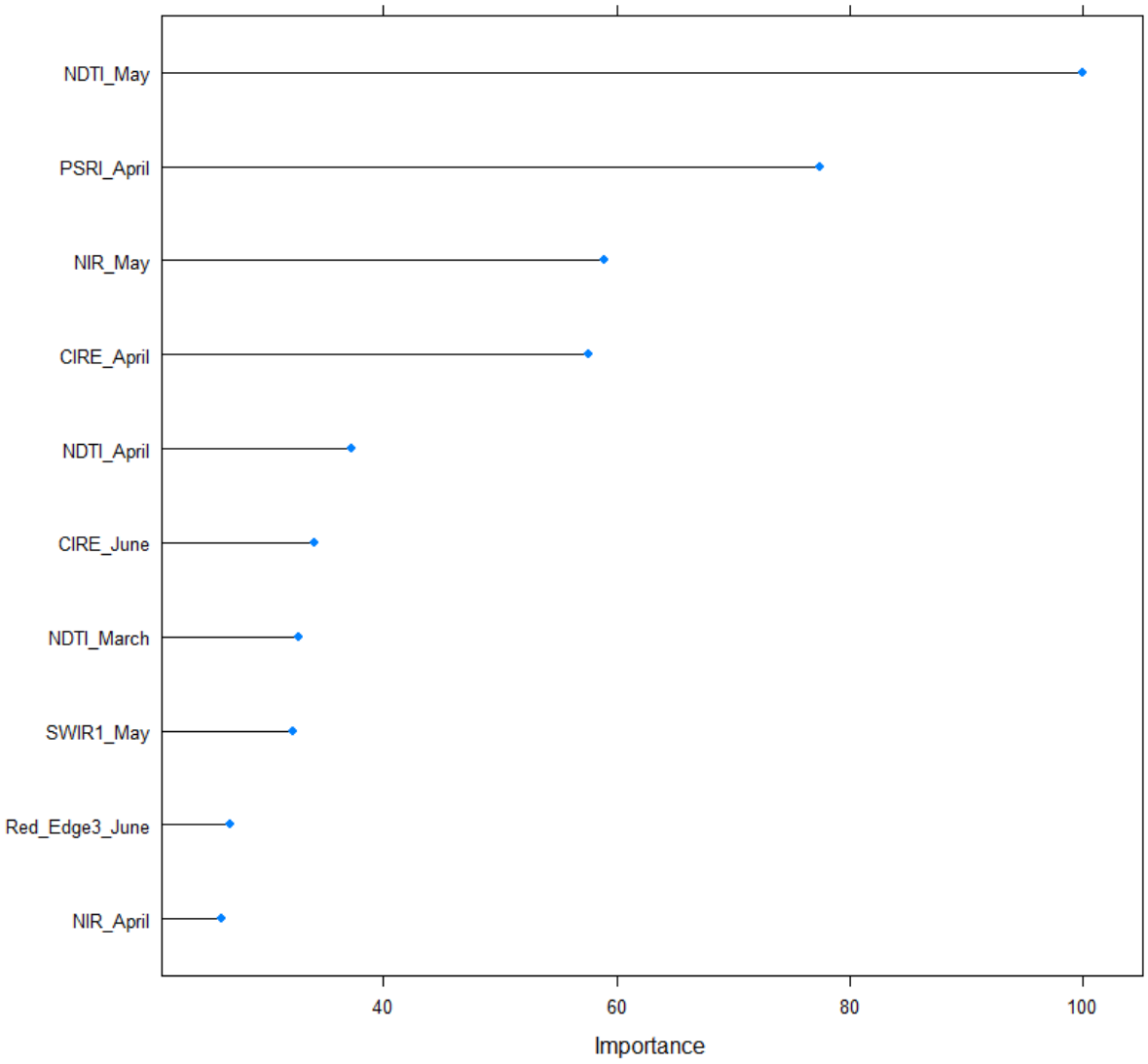

Figure S3. Variable Importance for SW region (HLS)

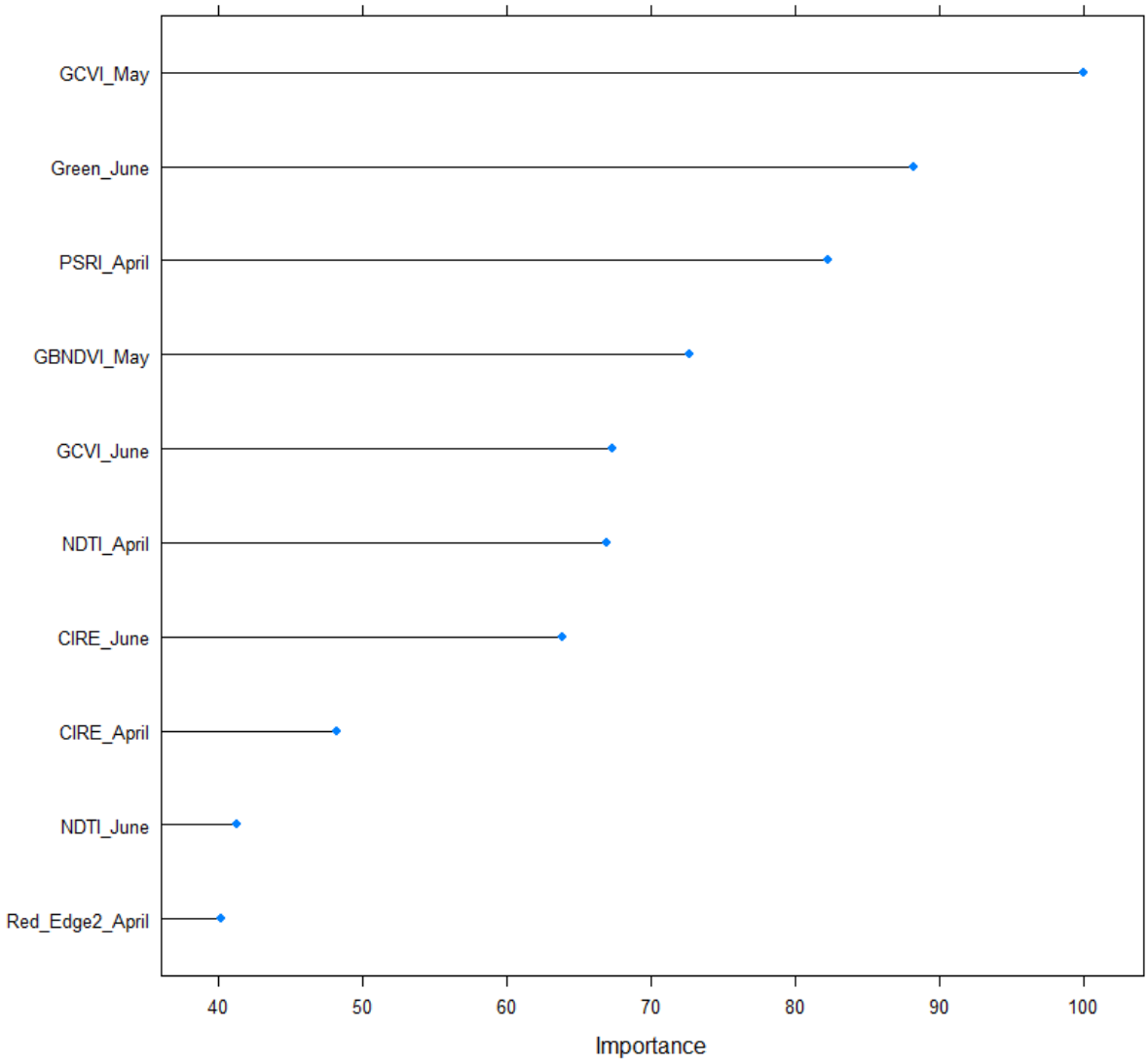

Figure S4. Variable Importance for SE region (HLS)