Received: 10 April 2019 | Revised: 10 October 2019 | Accepted: 16 October 2019
DOI: 10.1111/agec.12531

ORIGINAL ARTICLE

AGRICULTURAL ECONOMICS
The Journal of the International Association of Agricultural Economists
WILEY

Bringing automated, remote-sensed, machine learning methods to monitoring crop landscapes at scale

Xiaowei Jia[1] | Ankush Khandelwal[1] | David J. Mulla[2] | Philip G. Pardey[3] | Vipin Kumar[1]

[1]Department of Computer Science and Engineering, University of Minnesota, Twin Cities, Minneapolis, Minnesota

[2]Department of Soil, Water and Climate, University of Minnesota, Twin Cities, Minneapolis, Minnesota

[3]Department of Applied Economics, University of Minnesota, Twin Cities, Minneapolis, Minnesota

Correspondence
Vipin Kumar, Department of Computer Science and Engineering, University of Minnesota, Twin Cities, Minneapolis, MN 55455.
Email: kumar001@umn.edu

Funding information
NSF, Grant/Award Numbers: 1838159, 1739191, 1029711

Abstract
This article provides an overview of how recent advances in machine learning and the availability of data from earth observing satellites can dramatically improve our ability to automatically map croplands over long periods and over large regions. It discusses three applications in the domain of crop monitoring where machine learning (ML) approaches are beginning to show great promise. For each application, it highlights machine learning challenges, proposed approaches, and recent results. The article concludes with discussion of major challenges that need to be addressed before ML approaches will reach their full potential for this problem of great societal relevance.

KEYWORDS
deep learning, machine learning, monitoring crop landscapes, remote sensing

JEL CLASSIFICATION
N5, Q1

1 | INTRODUCTION

Agriculture is one of the most crucial ecosystems for human sustenance. Agricultural resources are witnessing tremendous supply-side stresses as a result of rapidly increasing population, suboptimal farming practices, increased pest damage occurrences due to climate change, and loss of productive land caused by human activities such as urbanization (Bebber, Holmes, & Gurr, 2014; d'Amour et al., 2017; Ortiz et al., 2008). Hence, multiagency international initiatives such as GEO-GLAM (Singh Parihar et al., 2012) and Global Yield Gap Atlas (Global Yield Gap Atlas, 2018) have been created with the goal of improving our ability to produce and share relevant, timely, and accurate trends and forecasts of crop productivity globally. Efforts within these initiatives have historically depended on surveys and self-reported statistics at regional and national scale, and thus are necessarily limited in scope and resolution both in space and time.

Advances in earth observation technologies have led to the acquisition of vast amounts of earth system data that can be used for monitoring changes on a global scale. In particular, a wide variety of instruments and sensors onboard satellites operated by the United States and other international agencies collect petabytes of data on a regular basis. For example, moderate resolution imaging spectroradiometer (MODIS) sensors onboard Terra and Aqua satellites have been collecting optical data daily at 500-m spatial resolution since 2000. Another commonly used dataset is from the Landsat series of satellites at biweekly temporal scale and 30-m spatial resolution since 1974. More recently, the Sentinel series of satellites launched by the European Space Agency have been capturing data both through optical and radar sensors with a spatial resolution of at least 10 m and temporal resolution of 5–10 days. The advent of microsatellite constellations for use in earth observation, such as those offered by Planet, is going to further increase the depth and breadth of remotely sensed data assets. These rich datasets hold great potential for cropland monitoring as they contain rich temporal and spectral information to map different types of crops and estimate the changes in the spatial distribution of different crops around the world due to climate, market, and policy changes (Alston, Babcock, & Pardey, 2010; Beddow & Pardey, 2015).

Machine learning models, which have found tremendous success in commercial applications, for example, computer vision and natural language processing (Liu et al., 2017), are increasingly being considered as alternatives (or at least complements) to surveys and self-reported data to provide accurate and timely information about crop locations and productivity on a global scale. In this article, we aim to provide an overview of how recent advances in machine learning approaches and the rapidly expanding availability of data from earth observing satellites can dramatically improve our ability to automatically map cropland change over long periods and large regions.

The remainder of this article is organized as follows. Sections 2, 3, and, 4 provide a discussion of three applications where machine learning approaches are beginning to show great promise. For each application, we highlight the machine learning challenges and then discuss proposed approaches and results. Section 5 concludes with a discussion of the major challenges that still need to be addressed before machine learning approaches will reach their full potential and enable the timely and reliable measurement of cropping agriculture to scale.

2 | CROPLAND MAPPING

Effective cropland mapping is important since it provides accurate and timely agricultural information and also helps track the consumption and transfer of water, nutrients, and energy. In this task, we aim to use machine learning techniques to detect what types of crops are planted at each location in a large target region.

Satellites are able to measure reflectance from bare soil, crop residue carried over to spring from fall harvest, and the growth and development of crops planted in spring. Croplands gradually turn from brown to green as crops emerge and grow. As crops senesce and are harvested, croplands turn from green to brown again. Crops differ in their planting dates and growth patterns, for example, some crops turn green much faster than others, and other crops stay green much later than other crops. The availability of multi-temporal data allows differentiation between crop types by modeling differences in greenness associated with the temporal growth processes of crops.

There are several challenges in applying machine learning models to map croplands. First, different crops are very likely to be confused with each other using the data captured on a single date. An individual remote sensing image can also contain much noise due to aerosols and acquisition errors (Karpatne, Jiang, Vatsavai, Shekhar, & Kumar, 2016). Although many researchers have successfully applied machine learning models to map croplands using a single remote-sensed image (Chen, Nasrabadi, & Tran, 2011; Jia et al., 2016), this approach hinges on the availability of imagery at a desired period during which the target land cover can be differentiated. Also, this approach cannot be generalized to more complex scenarios where each crop type can only be differentiated at a different point in time. Generally, crops can be best distinguished based on their cumulative growing patterns over time. For example, previous research (Sakamoto et al., 2010) shows that remote-sensed corn and soybean data look very similar for most dates of each year, but are differentiable at certain growth stages using their temporal patterns of phenologically driven plant development.

Moreover, while multi-temporal remote-sensed data cover the entire year at regular time intervals, crops show their distinctive temporal patterns only during certain period. These periods when crops can be better identified are also referred to as the "discriminative period." Therefore, it is critical to develop a classification model that automatically pays particular attention to the discriminative periods while reducing the potentially confounding impact of remote-sensed data from other periods.

As shown in previous research (Sakamoto et al., 2010), temporal features that capture differences in the growth patterns of plants are helpful in distinguishing among different crops. While traditional machine learning models utilize the feature-based or distance-based methods to classify multi-temporal data (Xing, Pei, & Keogh, 2010), recent advances in deep learning have provided much more effective ways of modeling complex temporal dependencies, which greatly improves the classification performance.

In our research, we have developed deep learning approaches to automatically extract such temporal features from multi-temporal remote sensing data. Among all the deep learning models, the recurrent neural networks (RNN) model is mostly widely used to handle temporal data. However, RNN is known to suffer from a vanishing gradient problem (Bengio, Simard, & Frasconi, 1994). Basically, this approach gradually loses the connection to previous time steps as time progresses. To address this limitation of RNNs, we use the long short-term memory (LSTM) model, which is an extension of the standard RNN model and has found tremendous success in handling sequential data in commercial applications (Jia, Wang, Khandelwal, Karpatne, & Kumar, 2019; Luong, Pham, & Manning, 2015; Salehinejad, Sankar, Barfett, Colak, & Valaee, 2017). This model can automatically extract complex long-term data dependencies from multi-temporal data and encode the extracted crop patterns into high-level feature representations.

Specifically, the LSTM model generates a high-level abstract temporal feature representation h_t at each time step t. This representation encodes not only the information at current time step, but also the inherited temporal information from previous time steps. The hidden representation is computed in a recurrent process as $h_t = LSTM(x_t, h_{t-1})$, where the function $LSTM(\cdot)$ is defined based on the structure of the LSTM model. Prior works in the context of commercial and healthcare applications have shown that using the extracted temporal features h_t for classification can achieve a much better performance than directly using the original input data (Goodfellow, Bengio, & Courville, 2016).

Next, we detect the discriminative period and use the extracted temporal features using LSTM to conduct the crop classification. For this task, we use an attention model (Luong et al., 2015) to identify the period after seeding and before harvesting since the crop characteristics can only be captured during this period. The attention model is widely used in machine translation and image captioning (Luong et al., 2015; You, Jin, Wang, Fang, & Luo, 2016) for its ability to automatically find a specific portion of input data that is relevant to the target output. In particular, the attention model produces an attention weight α_t for each time step t to measure its contribution to crop classification. Higher values for α_t indicate that the time step t contains more relevant information for identifying crops.

Next, the temporal features from all the time steps are aggregated based on the obtained attention weights, as $\sum_t \alpha_t h_t$. The aggregated features contain more information from the discriminative time steps given their higher attention weights. The aggregated features are then used to conduct crop classification by a fully connected layer (Jia, Li, Khandelwal et al., 2019).

This method was implemented to classify corn versus soybean pixels in southwestern Minnesota, United States, in 2016. Our evaluation focused on these two major crops since they account for over 90% area in southwestern Minnesota. More importantly, their labels provided by United States Department of Agriculture Cropland Data Layer (USDA CDL, n.d.) product are more accurate than other minor crop types and thus can better serve as the ground-truth in the evaluation. We used the MODIS MODI09A1 product (NASA Earthdata, 2019) as input features. This dataset provides global data for every 8 days at 500-m spatial resolution. At each date, MODIS dataset provides reflectance values on seven spectral bands for every location. To better learn short-term temporal patterns, we concatenate spectral features into a rolling 32-day window as a time step and slide the window by 8 days. Totally this generates 43 time steps in a year.

TABLE 1 The performance of different machine learning models in mapping corn versus soybean in 2016, 2015, and 2011

	2016-Test		2015-Test		2011-Test	
	AUC	F1	AUC	F1	AUC	F1
ANN	0.863	0.797	0.665	0.679	0.664	0.602
RF	0.863	0.788	0.672	0.688	0.662	0.523
SVM^{HMM}	0.813	0.790	0.715	0.698	0.706	0.688
1-NN^{DTW}	0.814	0.792	0.703	0.687	0.700	0.685
S2V	0.837	0.806	0.736	0.712	0.730	0.706
LSTM	0.865	0.807	0.767	0.718	0.762	0.704
LSTM^{ATT}	**0.909**	**0.811**	0.799	0.753	0.779	0.721
DA	**0.909**	**0.811**	**0.840**	**0.775**	**0.831**	**0.758**

Note. The training data are taken only from 2016. The performance is measured using area under the curve (AUC) score and F-1 measure. The values in bold represent the best performance in each column.

We randomly selected 1,000 data points (500 corn and 500 soybean) from different locations (i.e., different MODIS pixels) in southwestern Minnesota. We predicted the crop types for another set of 2,000 randomly selected locations (1,000 corn and 1,000 soybean) from southwestern Minnesota in 2016. We name our proposed method LSTM^{ATT} and compare it against multiple widely used machine learning baselines, including artificial neural networks (ANN), random forest (RF), structural support vector machines for sequence tagging (SVM^{HMM}) (Altun, Tsochantaridis, & Hofmann, 2003), the nearest neighbor classifier with dynamic time warping (1-NN^{DTW}) (Nayak, Mithal, Jia, & Kumar, 2018), sequence to vector (S2V) (Xun, Jia, & Zhang, 2016), and standard LSTM. Among these baseline methods, ANN and RF are directly applied to the concatenation of multi-temporal data and thus ignore the dependencies between different dates. The remaining baselines are all commonly used sequential models. The performance is reported in Table 1 (2016-test). The method domain adaptation (DA) in Table 1 will be introduced in Section 2.2.

According to the 2016-test in Table 1, our proposed method LSTM^{ATT} outperforms all other methods by a considerable margin. The comparison between LSTM and static baselines (ANN and RF) shows that the modeling of the crop's temporal profile can help detect land cover with considerably increased accuracy. LSTM also outperforms other sequential baselines (e.g., SVM^{HMM}, 1-NN^{DTW}, S2V) because LSTM can extract representative temporal patterns by exploring complex dependencies across different spectral bands and across time. Also, the improvement from LSTM to LSTM^{ATT} shows that the attention model assists in further improving the classification performance by explicitly modeling the discriminative period.

Our method LSTM^{ATT} detects the discriminative period (i.e., the period with highest attention weights) is from June

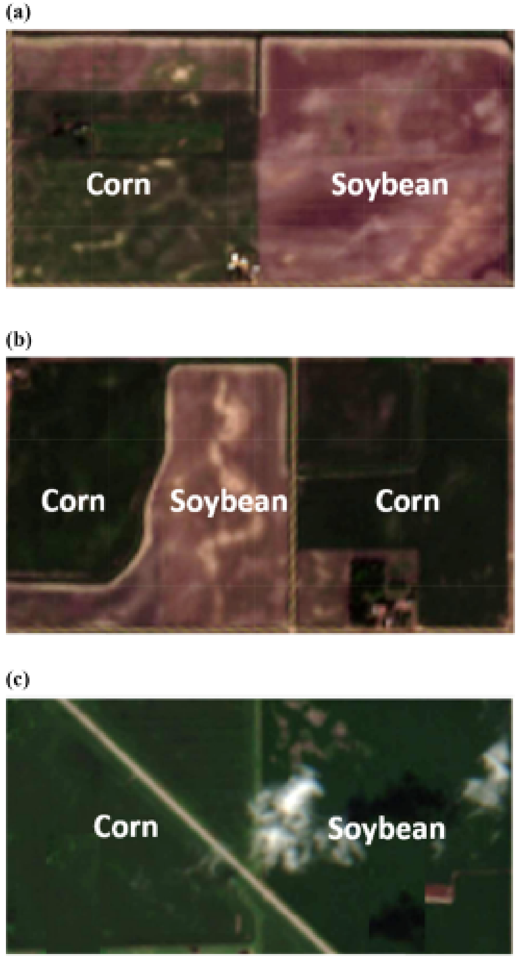

FIGURE 1 Sentinel-2 satellite images in RGB at 10-m spatial resolution. (a and b) Cropland patches with corn and soybean on June 24, 2016. Corn shows higher greenness level than soybean on this date. (c) Another cropland patch captured on August 6, 2016, where corn and soybean cannot be easily distinguished. Each image is approximately a 1500 m × 750 m area [Color figure can be viewed at wileyonlinelibrary.com]

9 to July 11. To verify this result, we drew on high-resolution Sentinel-2 images at 10-m resolution. Figure 1a,b shows some corn and soybean patches in four example regions using Sentinel-2 images on June 23, which show that corn patches turn green at a faster rate than nearby soybean patches. Such difference can also be verified by the normalized difference vegetation index (NDVI) time series as shown in Figure 2, where corn shows a much higher NDVI than soybeans at the early stage (red dashed box). It is easy to see from Sentinel images and NDVI series that these two crop types can be easily distinguished in this period.

Our method also detects the discriminative period (with higher attention weights than average) from July 19 to August 20. During this period, both corn and soybean samples show very high levels of greenness, making it difficult to distinguish between them (e.g., the August 6 Sentinel-2 image shown in Figure 1c). To verify that this period is indeed a discriminative period, we only use the multispectral features from July 19 to August 20 to train and test a simple ANN model, which produces area under the curve (AUC) and F1 score of 0.894 and 0.806, respectively. It is noteworthy this is better than the ANN baseline that is trained using full-year sequences (AUC, 0.863; F1, 0.797). This improvement demonstrates that our framework has potential to detect the discriminative period from the full multispectrum, which cannot even be observed directly by human experts.

Successful cropland mapping also helps estimate the distribution of winter cover crops that are planted in the autumn after harvest of the summer grain crop. Monitoring the spatiotemporal distribution of cover crops is important since they allow for a more effective use of natural and applied cropping inputs that both maintain or improve crop productivity while also enhancing the quality of the environment (Dabney, Delgado, & Reeves, 2001; Strock, Porter, & Russelle, 2004). In addition, increasing the diversity of vegetation on the landscape can provide numerous benefits including improved weed, pest, and disease resistance; improved soil water holding capacity and nutrient cycling; increased pollination services; and enhanced wildlife habitat (Derksen, Anderson, Blackshaw, & Maxwell, 2002; Kremen & Miles, 2012; Lin, 2011).

The rate of adoption of cover crops in the United States is increasing but remains small, in part, because they provide limited economic benefits to farmers. The proposed method estimates cover crop emergence and determines into which primary crop they were planted. This information provides useful insights for government and other (private and nonprofit) parties to enact new policies and practices that

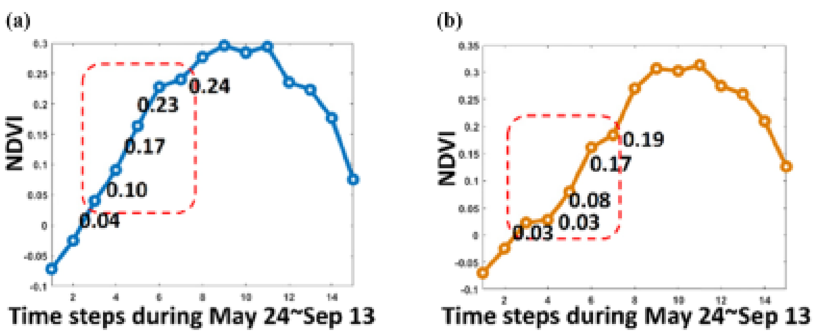

FIGURE 2 The average NDVI (greenness) series for (a) corn and (b) soybean during May 24–September 13, 2016. The red boxes indicate the detected discriminative period June 9–July 11, 2016 [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 2 The distribution of different crops and subsequent percentage and area for germination of cover crops in our study region (southwestern Minnesota)

	Total (Acre)	Cover crop (%)	Cover crop (Acre)
Corn	1,014,653	0.45	4,597
Soybean	716,215	0.48	3,421
Sweet corn	35,617	7.18	2,556
Spring wheat	9,229	45.19	4,171
Rye	561	62.17	349
Oats	523	40.85	214
Sugar beets	71,388	3.53	2,519
Dry beans	14,774	10.23	1,511
Peas	6,255	56.74	3,549
Total	1,869,215	1.22	22,888

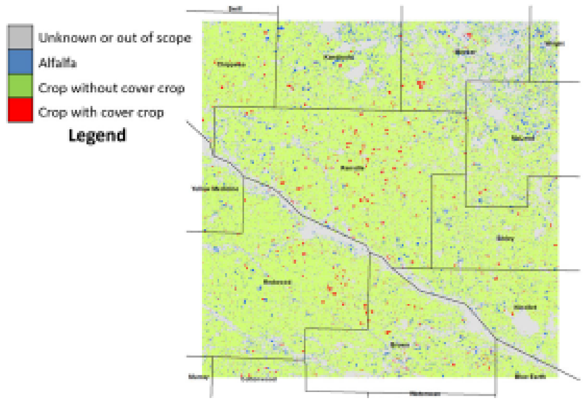

FIGURE 3 Fall 2016 map showing locations where cover crops have germinated and alfalfa has been planted in our study region. This area is 2,471,054 acre, or 10^8 Sentinel pixels (at 10-m resolution) [Color figure can be viewed at wileyonlinelibrary.com]

encourage planting cover crops in targeted areas that are especially vulnerable to soil and environmental degradation.

The widely used USDA CDL product does not explicitly label cover crops but does label the primary crops that are planted before cover crops emerge. Alternatively, the CDL may mistakenly classify cover crops as perennial crops such as alfalfa that are also green in early spring and late fall. Successful detection of cover crops requires analyzing the temporal profile of vegetation greenness in late fall and the transition patterns between primary crop types (e.g., corn and soybean) and cover crops. The most important feature for cover crops is that they are green after the primary crops are harvested. Successful detection also requires excluding other crops with green vegetation in late fall such as alfalfa. Alfalfa is green throughout the entire year, whereas cover crops are only green from fall (after harvest of primary crop) through the next spring (prior to planting of the next primary crop).

We show preliminary results of cover crop detection using Sentinel-2A imagery in 2016 for an area of 2.47 million acres in southwestern Minnesota, United States. In Table 2, we report the statistics for cover crop planting after nine primary crop types throughout the region. We observe that the uptake of cover crops is still limited for major crop types (0.45% of the cropped area for corn and 0.48% for soybean), and overall adoption is low at about 1% of the overall cropland area. In contrast, cover crops are frequently planted into minor crops such as spring wheat, rye, and peas that are harvested in August or early September, leaving adequate time for planting and germination of cover crops. In our implementation, we find that the labeling of alfalfa is critical for accurately assessing the ratio of the adoption of cover crops for each primary crop since cover crops are rarely planted and mistakenly classifying alfalfa can result in large variation of the result. In Figure 3, we show the locations where cover crop use has occurred, and differentiate these from locations where alfalfa has been planted.

3 | MAPPING CROP LANDSCAPES AT SCALE

Due to the high variability in the spectral properties of different crop cover types among regions and over time, it can be challenging to build machine learning models that perform well across multiple geographies, and time periods. For example, crops can be planted under different soil types, precipitation, and other weather conditions for different places and different years. Standard machine learning models only find the statistical relationships that fit the available training data. Therefore, a classification model learned from a specific year or a specific year cannot be generalized to other regions and time periods.

Generalizing learning models so they can perform well in different weather conditions makes it much easier to monitor croplands globally by avoiding the need to retrain models for places or years when ground-truth labels are not available or noisy. Furthermore, such techniques can potentially help understand the changes in crop-growing process across years.

Applying the LSTM^{ATT} method outlined in Section 2 to map crop landscapes at scale poses two performance problems. First, the classification performance will decrease since the statistical relationships learned from training data are unlikely to fit the spectral data generated under different growing conditions. Second, the detected discriminative period becomes less accurate. The parameters in the attention model are still estimated using training data and thus the generated attention weights in testing data can be less accurate in indicating the importance of each time step.

To mitigate this issue, an ideal strategy would be to train different models for each scenario independently. However, this is challenging in many real-world applications because of a paucity of labeled (ground-truthed) data, which is often only available for certain cropping regions and years.

In this work, we handle data heterogeneity through transfer learning methods (Pan & Yang, 2010) and especially DA approaches (Jiang, 2008) that are extensively used in computer vision and related applications for handling data heterogeneity. DA is a special case of transfer learning where different domains have the same feature space and class categories, but the joint probability distributions are not the same, such that $P_S(X,Y) \neq P_T(X,Y)$ between the source domain S and the target domain T. This situation is most common in earth observation data where the same datasets are available at a global scale, but the probability distributions of the data vary due to spatial and temporal heterogeneity. Here the source domain denotes the specific regions or years with sufficient labeled data, while the target domain denotes other target regions or years of our monitoring objective.

Deep learning-based DA methods greatly outperform previous methods for DA (Ganin et al., 2016; Tzeng, Hoffman, Saenko, & Darrell, 2017). The main reason behind their success is their ability to extract task specific features which that are also consistent across different domains (Li, Wang, Shi, Hou, & Liu, 2018; Venkateswara, Chakraborty, & Panchanathan, 2017; Williams, 2017). Due to the data heterogeneity and paucity of labels, researchers have also applied such techniques in remote sensing applications (Tuia, Persello, & Bruzzone, 2016).

However, these approaches mostly focus on individual image snapshots. In contrast, several approaches have been proposed for health-care data that explore the information transfer between multi-temporal data using RNN and its variants (Aswolinskiy & Hammer, 2017; Purushotham, Carvalho, Nilanon, & Liu, 2016; Yang, Salakhutdinov, & Cohen, 2017). All these methods leverage all the time steps equally in recurrent models to connect different domains, and thus lack the ability to avoid the transfer of non-informative time steps. Consequently, these approaches can be adversely affected by the variability in the non-informative period.

Here, we discuss an approach that handles this issue by combining the discriminative period detection technique (described in Section 2) with the DA approach. As mentioned in Section 2, in many situations, the complete temporal duration is not required to discriminate between different categories of interest. Since farmers can and do switch crop types across years, adaptation on the discriminative period is especially helpful for cropland mapping at scale because the model will not have to adapt the high variability before crops are planted.

The goal of DA is to learn a mapping from the target domain to the source domain, $g : T \rightarrow S$. This function aims to transform the data in a target domain to a distribution similar to the source domain such that the learned model can be applied to the transformed samples $g(x_T)$, where x_T are the samples from the target domain. To learn the mapping between the source domain and the target domain, a standard approach is to minimize the divergence between the hidden representation of two domains. Learning strategies such as adversarial deep learning have been used to minimize the divergence (Purushotham et al., 2016), which enforces that the hidden representation of the source domain cannot be distinguished from the target domain after applying the transformation $g(\cdot)$. Note that these hidden representations can potentially be extracted using the modeling framework introduced in Section 2.1.

Instead of directly conducting adversarial regularization on multispectral features or the hidden representation, which have been widely adopted in previous works (e.g., Purushotham et al., 2016), we apply the adversarial training on the weighted summation of hidden representation at different time steps to take account of seasonality in cropping patterns. In this way, the adaptation process can pay more attention to the discriminative period. Specifically, we utilize the weights obtained from the attention model described in Section 2.1, as these weights indicate the importance of each time step in classification.

However, previous research has shown that the attention model can be severely impacted when applied across different domains (Kang, Zheng, Yan, & Yang, 2018). Hence, to utilize the attention model under different weather conditions, it is important to ensure the robustness of the attention model when applied across different domains. To address this problem, we introduce another regularizer on the difference of attention weights between the original samples x_T and the transformed samples $g(x_T)$. This mechanism has shown to improve the robustness of the attention model and successfully fix the attention weights.

We implement the proposed DA method on the same training dataset with Section 2.1, which was collected in 2016. Here, we test the model using two sets of data points collected in 2015 and 2011. Each test dataset contains 2,000 locations that are randomly selected from southwestern Minnesota.

According to Table 1, the performance of each method is in general degraded in 2015-test and 2011-test compared with the tests in the same growing condition (2016-test). The DA-based approach shows superior performance compared with LSTM^{ATT} in 2015 and 2011 since it can reduce the divergence between source and target domains.

Now we assess the impact on the attention model arising from data heterogeneity. Figure 4a shows the obtained attention weights for the corn locations by LSTM^{ATT} in 2016 (Train) and 2015 (Test), as well as the obtained relevance scores by DA in 2015. Figure 4b shows the obtained attention weights for the corn locations by LSTM^{ATT} in 2016 (Train) and 2011 (Test), as well as the obtained relevance scores by DA in 2011.

For both tests, we can observe that the LSTM^{ATT} networks cannot detect a period with obviously higher attention weights when it is directly applied to the testing scenario. Therefore, it

AGRICULTURAL ECONOMICS
The Journal of the International Association of Agricultural Economists
WILEY

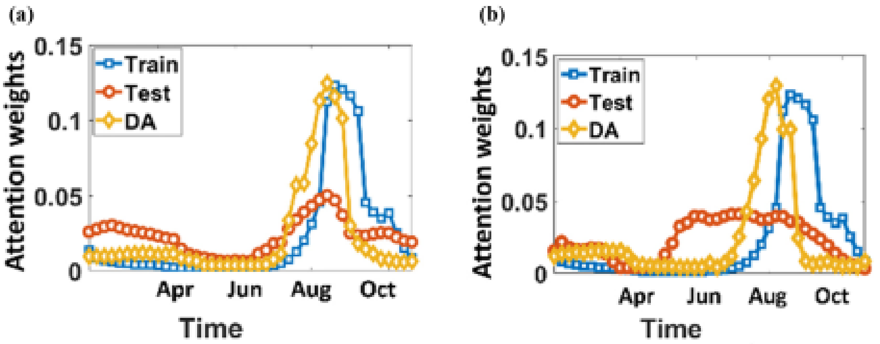

FIGURE 4 The impact of heterogeneity on the attention model in (a) 2015-test, and (b) 2011-test for cropland. Train: the attention weights on training data in 2016. Test: the attention weights on test data by directly applying the LSTM^{ATT} model. DA: the obtained attention weights on test data using the proposed DA method [Color figure can be viewed at wileyonlinelibrary.com]

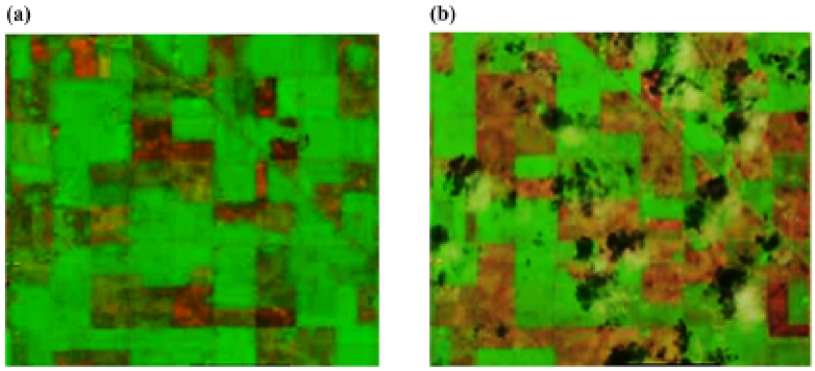

FIGURE 5 The Landsat images (in RGB, 30-m resolution) for an example region in southwestern Minnesota in (a) 2016 and (b) 2015 at the beginning of July [Color figure can be viewed at wileyonlinelibrary.com]

cannot precisely capture the discriminative period. In contrast, DA is capable of mitigating the impact of variability across domains and thus producing meaningful relevance scores.

The proposed method also enables interpreting the shift of discriminative period across years. For example, we can easily observe from Figure 4a that the crops in 2015 are planted earlier than the crops in 2016. This monitoring capacity is important in tracking variation in farmer behavior in correspondence to weather changes and other, perhaps economic, factors. To verify this finding, we show high-resolution Landsat images at the beginning of July in 2015 (Figure 5a) and 2016 (Figure 5b). It can be seen that the selected region shows higher greenness level at this selected time in 2015 than in 2016. Hence, farmers are more likely to grow crops earlier in 2015.

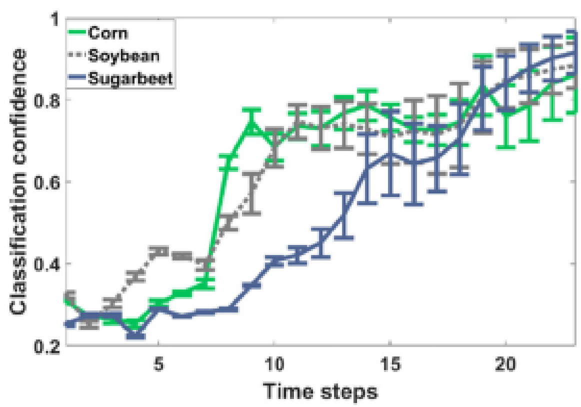

FIGURE 6 Confidence progression for corn, soybean, and sugar beets from April 5 to November 11 in 2016 (totally 23 time steps). The error bar represents the standard deviation [Color figure can be viewed at wileyonlinelibrary.com]

4 | EARLY CROP DETECTION

Monitoring crop types and planting area is timely science that can inform a myriad of choices by a host of public and private (including farmer) decision makers with consequential outcomes for cropping agriculture worldwide and related environmental concerns. Hence, it is of great interest to governments and companies alike to obtain relevant crop related information in the current year as soon as possible.

Traditional RNN-based classification methods generate class labels only for the entire sequence (Jia, Li, Khandelwal et al., 2019; Nayak et al., 2018) or for each time step separately (Jia et al., 2017; Zhang et al., 2017). However, these methods cannot be used to conduct classification at an early stage.

We introduce a new method to identify crops at an early stage (Jia, Wang et al., 2019). Preliminary results in detecting the spatial landscape of corn, soybean, and sugar beets are shown in Figure 6. Our method identifies that corn-related pixels quickly gains confidence at the eighth and ninth time steps, which correspond to June 14–24. To validate the correctness of this finding, we show the red, green, and blue (RGB) image of an example region captured on June 24 in Figure 7a. We can clearly see that in the early growing season,

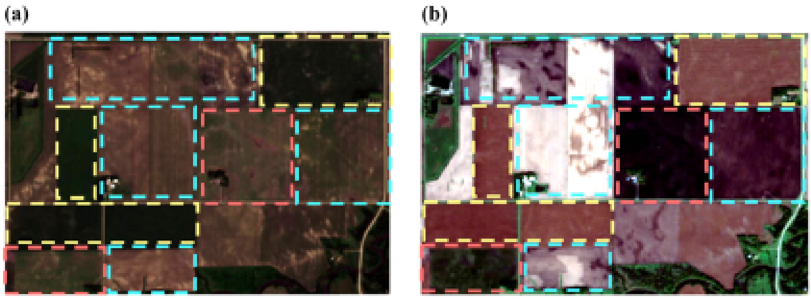

FIGURE 7 The Sentinel imagery in RBG captured on (a) June 24 and (b) October 5. Color legend for blocks: yellow, corn; blue, soybean; red, sugar beet. Each area is approximately a 3000 m × 2000 m area [Color figure can be viewed at wileyonlinelibrary.com]

corn turns into green more quickly than soybean and sugar beets, and therefore can be identified in this period.

We can also observe that the sugar beets samples still gain confidence after October. We show another RGB image on October 5 (the same example region) in Figure 7b. While corn and soybeans have been harvested, the cropland of sugar beets still remains green. These results show promising potential for applying machine learning approaches to track classification confidence over time. The progression of classification confidence can potentially lead to an early detection of specific crop types when they reach a sufficient (actionable) confidence level.

5 | CONCLUSION

As demonstrated by the results in this article, recent advances in machine learning and the availability of earth-observing satellite data can greatly improve our capability to monitor croplands over space and over time. While this technology is poised to play a key role in addressing issues related to food security at global scale, advances are still needed in many areas to realize this goal.

For example, while machine learning techniques are beginning to show success in extracting temporal patterns to create maps of crops at the pixel level, these approaches (Jia et al., 2017; Jia, Li, Khandelwal et al., 2019; Lyu, Lu, & Mou, 2016) have only been designed for remote sensing data that is available with high frequency, for example, MODIS (250 m, daily). However, the spatial resolution of such data is quite poor, which makes them unsuitable for monitoring small-scale farms that are quite common in many parts of the world. Higher spatial resolution data (at 10-m resolution) is available from Sentinel, but these images are only available every 5 or 10 days (depending on specific years and regions). Hence new advanced machine learning models need to be developed that combine coarse-resolution data and high-resolution data—for example, Sentinel (10 m, every 10 days in 2016) and Landsat (30 m, every 16 days)—such that it can leverage the temporal patterns from low-resolution data while also mapping croplands at high resolution.

Finally, traditional machine learning and data mining algorithms fail to take advantage of the wealth of information about physical principles or human behaviors and practices that govern or have huge impact on the crop growth. For example, a unique aspect of crop yield estimation that differentiates it from standard classification or regression tasks in machine learning is that crop-growing processes under specific environments are governed by relevant physical principles. Researchers in the agricultural community have also built biologically based (or crop process-based) models to simulate these physical principles (Jones et al., 2003; Srinivasan, Zhang, & Arnold, 2010). The ability to integrate such mechanistic knowledge into a machine learning framework will be key to advancing the state of the art in estimating crop yields.

ACKNOWLEDGMENT

This work was funded by the NSF awards 1838159, 1739191, and 1029711. Access to computing facilities was provided by the Minnesota Supercomputing Institute.

REFERENCES

Alston, J. M., Babcock, B. A., & Pardey, P. G. (2010). *The shifting patterns of agricultural production and productivity worldwide*. Ames, IA: Midwest Agribusiness Trade Research and Information Center.

Altun, Y., Tsochantaridis, I., & Hofmann, T. (2003). Hidden Markov support vector machines. In T. Fawcett & N. Mishra (Eds.), *Proceedings of the Twentieth International Conference on Machine Learning (ICML-03)* (pp. 3–10). Palo Alto, CA: AAAI Press.

Aswolinskiy, W., & Hammer, B. (2017). Unsupervised transfer learning for time series via self-predictive modelling–First results. *Proceedings of the Workshop on New Challenges in Neural Computation (NC2)*, Machine Learning Reports (Vol. 3). Bielefeld, Germany: Universität Bielefeld, CITEC.

Bebber, D. P., Holmes, T., & Gurr, S. J. (2014). The global spread of crop pests and pathogens. *Global Ecology and Biogeography*, *23*(12), 1398–1407.

Beddow, J. M., & Pardey, P. G. (2015). Moving matters: The effect of location on crop production. *Journal of Economic History*, *75*(1), 219–249.

Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, *5*(2), 157–166.

Chen, Y., Nasrabadi, N. M., & Tran, T. D. (2011). Hyperspectral image classification using dictionary-based sparse representation. *IEEE*

Transactions on Geoscience and Remote Sensing, *49*(10), 3973–3985.

Dabney, S. M., Delgado, J. A., & Reeves, D. W. (2001). Using winter cover crops to improve soil and water quality. *Communications in Soil Science and Plant Analysis*, *32*(7–8), 1221–1250.

d'Amour, C. B., Reitsma, F., Baiocchi, G., Barthel, S., Güneralp, B., Erb, K. H., … Seto, K. C. (2017). Future urban land expansion and implications for global croplands. *Proceedings of the National Academy of Sciences, USA*, *114*(34), 8939–8944.

Derksen, D. A., Anderson, R. L., Blackshaw, R. E., & Maxwell, B. (2002). Weed dynamics and management strategies for cropping systems in the northern Great Plains. *Agronomy Journal*, *94*(2), 174–185.

Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., … Lempitsky, V. (2016). Domain-adversarial training of neural networks. *The Journal of Machine Learning Research*, *17*(1), 2096–2030.

Global Yield Gap Atlas (2018). Retrieved from http://www.yieldgap.org/

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. Cambridge, MA: MIT Press.

Jia, X., Khandelwal, A., Gerber, J., Carlson, K., West, P., & Kumar, V. (2016). Learning large-scale plantation mapping from imperfect annotators. In *2016 IEEE International Conference on Big Data (Big Data)* (pp. 1192–1201). Washington, DC: IEEE.

Jia, X., Khandelwal, A., Nayak, G., Gerber, J., Carlson, K., West, P., & Kumar, V. (2017). Incremental dual-memory LSTM in land cover prediction. In S. Matwin, S. Yu, & F. Farooq (Chairs), *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 867–876). New York, NY: ACM.

Jia, X., Li, S., Khandelwal, A., Nayak, G., Karpatne, A., & Kumar, V. (2019). Spatio context-aware networks for mining temporal discriminative period in land cover detection. In T. Berger-Wolf & N. Chawla (Eds.), *Proceedings of the 2019 SIAM International Conference on Data Mining* (pp. 513–521). Philadelphia, PA: Society for Industrial and Applied Mathematics.

Jia, X., Li, S., Zhao, H., Kim, S., & Kumar, V. (2019). Towards robust and discriminative sequential data learning: When and how to perform adversarial training? In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 1665–1673). New York, NY: ACM.

Jia, X., Wang, M., Khandelwal, A., Karpatne, A., & Kumar, V. (2019). Recurrent generative networks for multi-resolution satellite data: An application in cropland monitoring. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence* (pp. 2628–2634). AAAI Press.

Jiang, J. (2008). *A literature survey on domain adaptation of statistical classifiers*. *3*, 1–12. Retrieved from http://www.mysmu.edu/faculty/jingjiang/papers/da_survey.pdf

Jones, J. W., Hoogenboom, G., Porter, C. H., Boote, K. J., Batchelor, W. D., Hunt, L. A., … Ritchie, J. T. (2003). The DSSAT cropping system model. *European Journal of Agronomy*, *18*(3–4), 235–265.

Kang, G., Zheng, L., Yan, Y., & Yang, Y. (2018). Deep adversarial attention alignment for unsupervised domain adaptation: The benefit of target expectation maximization. *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 401–416).

Karpatne, A., Jiang, Z., Vatsavai, R. R., Shekhar, S., & Kumar, V. (2016). Monitoring land-cover changes: A machine-learning perspective. *IEEE Geoscience and Remote Sensing Magazine*, *4*(2), 8–21.

Kremen, C., & Miles, A. (2012). Ecosystem services in biologically diversified versus conventional farming systems: Benefits, externalities, and trade-offs. *Ecology and Society*, *17*(4).

Li, Y., Wang, N., Shi, J., Hou, X., & Liu, J. (2018). Adaptive Batch Normalization for practical domain adaptation. *Pattern Recognition*, *80*, 109–117.

Lin, B. B. (2011). Resilience in agriculture through crop diversification: Adaptive management for environmental change. *BioScience*, *61*(3), 183–193.

Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., & Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, *234*, 11–26.

Luong, M. T., Pham, H., & Manning, C. D. (2015). Effective approaches to attention-based neural machine translation. *arXiv*, 1508.04025.

Lyu, H., Lu, H., & Mou, L. (2016). Learning a transferable change rule from a recurrent neural network for land cover change detection. *Remote Sensing*, *8*(6), 506.

NASA Earthdata (2019). *MODIS product table*. Retrieved from https://lpdaac.usgs.gov/products/mod09a1v006/

Nayak, G., Mithal, V., Jia, X., & Kumar, V. (2018). Classifying multivariate time series by learning sequence-level discriminative patterns. In M. Ester & D. Pedreschi (Eds.), *Proceedings of the 2018 SIAM International Conference on Data Mining* (pp. 252–260). Philadelphia, PA: Society for Industrial and Applied Mathematics.

Ortiz, R., Sayre, K. D., Govaerts, B., Gupta, R., Subbarao, G. V., Ban, T., … Reynolds, M. (2008). Climate change: Can wheat beat the heat? *Agriculture, Ecosystems and Environment*, *126*(1–2), 46–58.

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, *22*(10), 1345–1359.

Purushotham, S., Carvalho, W., Nilanon, T., & Liu, Y. (2016). Variational recurrent adversarial deep domain adaptation. Retrieved from https://openreview.net/pdf?id=rk9eAFcxg

Sakamoto, T., Shibayama, M., Takada, E., Inoue, A., Morita, K., Takahashi, W., … Kimura, A. (2010). Detecting seasonal changes in crop community structure using day and night digital images. *Photogrammetric Engineering and Remote Sensing*, *76*(6), 713–726.

Salehinejad, H., Sankar, S., Barfett, J., Colak, E., & Valaee, S. (2017). Recent advances in recurrent neural networks. *arXiv*, 1801.01078.

Singh Parihar, J., Justice, C., Soares, J., Leo, O., Kosuth, P., Jarvis, I., … Becker-Reshef, I. (2012), July. *GEO-GLAM: A GEOSS-G20 initiative on global agricultural monitoring*. Paper presented at the 39th COSPAR Scientific Assembly, Mysore, Karnataka, India (Vol. 39, p. 1451).

Srinivasan, R., Zhang, X., & Arnold, J. (2010). SWAT ungauged: Hydrological budget and crop yield predictions in the Upper Mississippi River Basin. *Transactions of the ASABE*, *53*(5), 1533–1546.

Strock, J. S., Porter, P. M., & Russelle, M. P. (2004). Cover cropping to reduce nitrate loss through subsurface drainage in the northern US Corn Belt. *Journal of Environmental Quality*, *33*(3), 1010–1016.

Tuia, D., Persello, C., & Bruzzone, L. (2016). Domain adaptation for the classification of remote sensing data: An overview of recent advances. *IEEE Geoscience and Remote Sensing Magazine*, *4*(2), 41–57.

Tzeng, E., Hoffman, J., Saenko, K., & Darrell, T. (2017). Adversarial discriminative domain adaptation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 7167–7176). Washington, DC: IEEE.

USDA CDL. (n.d.). *CropScape–Cropland data layer*. Retrieved from https://nassgeodata.gmu.edu/CropScape/

Venkateswara, H., Chakraborty, S., & Panchanathan, S. (2017). Deep-learning systems for domain adaptation in computer vision: Learning transferable feature representations. *IEEE Signal Processing Magazine*, *34*(6), 117–129.

Williams, J. M. (2017). Deep learning and transfer learning in the classification of EEG signals. https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1153&context=computerscidiss

Xing, Z., Pei, J., & Keogh, E. (2010). A brief survey on sequence classification. *ACM SIGKDD Explorations Newsletter*, *12*(1), 40–48.

Xun, G., Jia, X., & Zhang, A. (2016). Detecting epileptic seizures with electroencephalogram via a context-learning model. *BMC Medical Informatics and Decision Making*, *16*(2), 70.

Yang, Z., Salakhutdinov, R., & Cohen, W. W. (2017). Transfer learning for sequence tagging with hierarchical recurrent networks. *arXiv*, 1703.06345.

You, Q., Jin, H., Wang, Z., Fang, C., & Luo, J. (2016). Image captioning with semantic attention. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 4651–4659). Washington, DC: IEEE.

Zhang, Y., Lin, C., Chi, M., Ivy, J., Capan, M., & Huddleston, J. M. (2017). LSTM for septic shock: Adding unreliable labels to reliable predictions. In *2017 IEEE International Conference on Big Data* (*Big Data*) (pp. 1233–1242). Washington, DC: IEEE.

How to cite this article: Jia X, Khandelwal A, Mulla DJ, Pardey PG, Kumar V. Bringing automated, remote-sensed, machine learning methods to monitoring crop landscapes at scale. *Agricultural Economics*. 2019;50:41–50. https://doi.org/10.1111/agec.12531

APPENDIX

We have released the remote-sensed satellite data for our study region in southwestern Minnesota, Unites States.[1] The released dataset contains the multispectral data collected by Sentinel-2A satellite (at 10-m resolution) and MODIS Terra satellite (at 500-m resolution). The TIF file "USDA_2016.tif" contains the crop labels provided by USDA CDL for each location in our study region at 30-m resolution.

[1] https://drive.google.com/drive/folders/14mpxMSeOFufwIxT7GQWcUUZFsO2zvi27?usp=sharing.