DYNAMIC TIME WARPING FOR CROPS MAPPING

M. Belgiu [1,*], Y. Zhou [1], M. Marshall [1], A. Stein [1]

[1] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, 7500 AE, Enschede, the Netherlands-
(m.belgiu, m.t.marshall,a.stein)@utwente.nl, zhouyibo2017@gmail.com

Commission III, WG III/10

KEY WORDS: Dynamic Time Warping, Multi-temporal satellite images, Mapping

ABSTRACT:

Dynamic Time Warping (DTW) has been successfully used for crops mapping due to its capability to achieve good classification results when a reduced number of training samples and irregular satellite image time series is available. Despite its recognized advantages, DTW does not account for the duration and seasonality of crops and local differences when assessing the similarity between two temporal sequences. In this study, we implemented a Weighted Derivative modification of DTW (WDDTW) and compared it with DTW and Time Weighted Dynamic Time Warping (TWDTW) for crops mapping. We show that WDDTW outperformed DTW achieving an overall accuracy of 67 %, whereas DTW obtained an accuracy of 57%. Yet, TWDTW performed better than both methods obtaining an accuracy of 88%.

1. INTRODUCTION

The earth's population exceeds 7 billion (Brelsford et al., 2018). This population growth puts pressure on the food supply systems across the globe. Therefore, food security related challenges are high on the agenda of the Sustainable Development Goals (SDG) (United-Nations, 2015), in particular SDG Goal 2-Zero Hunger. To be able to address these challenges, decision-makers require accurate and spatially explicit information on crops (Fritz et al. 2015). Such information is of paramount importance for evidence-based action and policy formulation. The importance of remote sensing data for generating this information is widely recognized.

Spatially and temporally explicit information on cropland extent and crop types is a prerequisite for monitoring the multiannual dynamics of crop production (Bégué et al. 2018) and associated water consumption (Meier et al. 2018). Research has been carried out in recent years to develop automated methods for crop mapping. Yan and Roy (2014) proposed an object-based method to delineate crop fields in multi-temporal satellite images available through Web Enabled Landsat Data project. Waldner et al (2015) described a cropland areas classification method that relies on expert knowledge to derive features that are expected to remain the same from one year to another. The efficiency and transferability of this method were tested in Argentina, Belgium, China, and the Ukraine. Salmon et al. (2015) mapped three classes of agricultural land use (rain-fed, irrigated, and paddy croplands) by means of machine learning techniques and by combining data from a number of different sources including MODIS imagery, agricultural census data, meteorological measurements, and surface moisture measurements. Xiong et al. (2017) created a cropland areas map of the African continent using the Google Earth Engine and a MODIS dataset. Dahal et al. (2018) produced a 250 m spatial resolution cropland map using MODIS NDVI, weather, and climate information, together with elevation data.

Important research has been carried out for crops mapping from multi-temporal satellite images using machine learning classifiers such as random forest (Wang et al., 2019), or deep learning (Zhong et al., 2019). These classifiers are, however, challenged by the following problems: 1) lack of a large number of training samples required for achieving good classification results; 2) large variability of crops phenology across geographic regions and time, which prevents the reusability of these classifiers from one region to another and from one year to another; 3) lack of a consistent multi-temporal satellite image stack to be processed due to the presence of clouds.

The increasing availability of high-resolution satellite image time series has created new premises for crops mapping. In an ideal scenario, we can compare two temporal sequences representing the growth pattern of target crops using simple metrics such as the Euclidean distance. In reality though, the temporal sequences for the crops of interest may not line up on the time axis because of the different climatic conditions or different agricultural practices. In this case, we need to warp the measurements on the time axis for a better alignment. Dynamic Time Warping (DTW) is an efficient solution for a non-linear alignment of two temporal sequences (Sakoe, Chiba, 1978).

DTW obtained good crops classification results with a reduced number of training samples and with multi-temporal images that present gaps caused by the presence of clouds (Belgiu, Csillik, 2018; Csillik et al., 2019; Guan et al., 2018; Maus et al., 2016; Petitjean, et al., 2011). DTW has initially been developed for speech recognition applications (Sakoe, Chiba, 1978), and it has been adopted in other domains including medicine (Tsevas, Iakovidis, 2010) or robotics (Johnen, Kuhlenkoetter, 2016). DTW is used to assess the similarity between the phenological patterns generated from, for example, Normalized Difference vegetation Index (NDVI, of the labelled samples with the phenological patterns computed for the remaining unlabelled

---

* Corresponding author

pixels. When aligning two phenological patterns, the following factors might affect the similarity assessment results:

(1) *noise*: the presence of clouds might introduce noise in the evaluated phenological patterns;

(2) *amplitude scaling*: two phenological patterns representing the same crop might have different amplitude, i.e. different NDVI values;

(3) *time-shifting*: the phenological pattern of the same crops might be shifted because of the difference in the time of cultivation;

(4) *time scaling*: one phenological pattern of a crop is shorter than the other one.

Despite its recognized advantages, DTW has some limitations for agriculture applications that need to be carefully addressed. First, DTW does not account for the crop growth phase, i.e. duration and seasonality (Maus et al., 2016). Second, DTW addresses the differences of the spectral measurements (e.g. the peak on the NDVI axis is higher on the test sequence than on the reference sequence) by warping the values on the time axis (Sakoe, Chiba, 1978).

The above-mentioned problems have been addressed by introducing different modifications of DTW. For example, Maus et al. (2016) proposed the so-called Time-Weighted Dynamic Time Warping (TWDTW) method that makes use of a time constraint to calculate the distance between two phenological patterns. Due to this time constraint, it accounts for the crop growing phase. TWDTW has successfully been applied for land cover classification in a tropical area (Maus et al., 2016). The authors reported an overall accuracy of 87.32%, as compared to 70% obtained by DTW. Belgiu, Csillik (2018) applied TWDTW for crops mapping from Sentinel-2-based NDVI layers in three agricultural landscapes from Romania, Italy and USA and obtained an overall accuracy of 96%. Guan et al. (2018) applied successfully an open-boundary locally weighted DTW distance method to NDVI time-series from MODIS for cropland mapping in Southeast Asia.

In contrast to DTW and TWDTW, Weighted Derivative Dynamic Time Warping (WDDTW) (Jeong et al., 2011) takes into account the first derivative of the evaluated sequences and penalizes the high phase differences. Jeong et al. (2011) compared WDDTW and DTW using time series datasets from different application domains and concluded that WDDTW outperformed DTW in all evaluated cases.

The goal of this paper is to implement and evaluate to what extent TWDTW and WDDTW can overcome the challenges encountered by DTW when classifying crops. It is the first study dedicated to implementing the weighted derivative variation of DTW for classifying crops from satellite image time series.

2. STUDY AREA AND DATASET

The study area is situated in Imperial County, southern California, USA. Seven crops were investigated, each of them occupying at least 3% of the study area: alfalfa, fallow, other hay, onion, sugar beet, winter wheat, and lettuce (Figure 1).

We used the red and near-infrared spectral bands (band 4 and band 8A with a spatial resolution of 10 m) acquired by Sentinel-2 MultiSpectral Instrument (MSI) to calculate the NDVI for eighteen images (September 2016-September 2017). These images were downloaded from Copernicus Open Access Hub. We used the atmospherically corrected satellite images as input in our study.
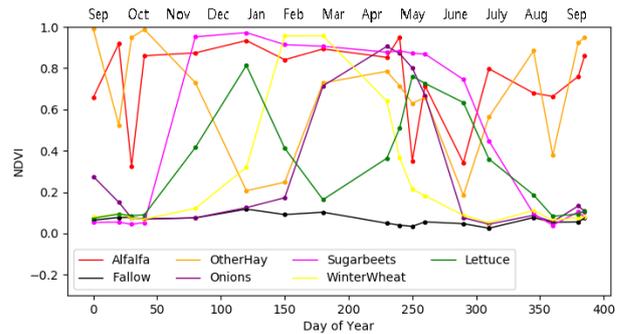

Figure 1. Phenological patterns of the target crops

3. METHODOLOGY

A three-step methodology was implemented in this study: (i) generation of NDVI-based temporal sequences; (ii) implementing DTW, TWDTW and WDDTW; (iii) evaluating the results using standard classification accuracy metrics.

3.1 Generating the NDVI stack

The NDVI generated from 10 m resolution Sentinel-2 red and near-infrared spectral bands was used to compute the temporal phenological patterns of the target classes:

$$NDVI = \frac{NIR - RED}{NIR + RED} \qquad (1)$$

NDVI was used in this study because of its efficiency for vegetation phenology studies (Yan, Roy, 2014).

3.2 Implementing and testing Dynamic Time Warping, Time Weighted Dynamic Time Warping, and Weighted Derivative Time Warping

DTW uses the Euclidean distance for similarity measurement between two temporal sequences. Given two time series sequences $\mathbf{U} = \{u_1, u_2, …, u_n\}$ and $\mathbf{V} = \{v_1, v_2, …, v_m\}$ with lengths n and m, a n-by-m matrix $\mathbf{D}_{base} = (d_{base}(u_i,v_j))_{n \times m}$ stores the Euclidean distances between $u_i \in \mathbf{U} \ \forall \ i = 1, 2, …, n$ and $v_j \in \mathbf{V} \ \forall \ j = 1, 2, … , m$.

$$d_{base(i,j)} = | u_i - v_j | . \qquad (2)$$

The DTW distance matrix D is obtained as the recursive sum of the minimum distances:

$$d_{i,j} = d_{base(i,j)} + \min\{d_{i-1,j}, d_{i-1,j-1}, d_{i,j-1}\} . \qquad (3)$$

TWDTW adopts a weight into the calculation of DTW distance. The weight is controlled by a Modified Logistic Weight Function (MLWF) which takes into account the phase difference between two matching points. For example, if a point in one sequence has a larger time lag than a point in another sequence, the weight between these two points will be larger

(Maus et al., 2016). The weight can be fine-tuned when measuring the similarity between two sequences. This can be done by defining different values of two MLWF parameters, i.e. half-length of the sequence and the level of penalization for the points with larger phase difference. Thus, TWDTW relies on a weight ω given to the d_{base},

$$d_{base(i,j)} = \omega_{i,j} \, | \, u_i - v_j \, | \qquad (4)$$

The WDTW distance wd is calculated as follows:

$$wd_{i,j} = \omega_{i,j} \, | \, u_i - v_j \, | + \min\{wd_{i-1,j}, wd_{i-1,j-1}, wd_{i,j-1}\} \qquad (5)$$

The weight ω is defined using a modified logistic weighted function:

$$\omega_{i,j} = \frac{1}{1 + e^{-\alpha(g(t_i, t_j) - \beta)}} \qquad (6)$$

WDDTW transforms the original temporal sequences generated using e.g. NDVI into higher level features containing the shape information of that sequence. The equation for transforming data point u_i in sequence U is presented below:

$$d_i^u = \frac{(u_i - u_{i-1}) + ((u_{i+1} - u_{i-1})/2)}{2}, 1 < i < n, \qquad (7)$$

where n = length of sequence U.

The weighted version of WDDTW is obtained as follows (Jeong et al., 2011):

$$wdd_{i,j} = \omega_{i,j} \, | \, d_i^u - d_j^v \, | + \min\{wdd_{i-1,j}, wdd_{i-1,j-1}, wdd_{i,j-1}\} \qquad (7)$$

where d_i^u and d_j^v are the transformed sequences from the sequences U and V with lengths n and m, respectively.

The three methods evaluated in this study have been implemented in Python and are available to interested users per request.

3.3 Training and validation samples

The training and validation samples used for this study have been generated using the CropScape - Cropland Data Layer (CDL) for 2017. CDL data are collected by the United States Department of Agriculture (USDA), National Agricultural Statistics Service (Boryan, 2011). Given that the spatial resolution of CropScape data is 30 m, we decided to resample the Sentinel-2 spectral bands from 10 m to 30 m. The spatial distribution of the target crops is depicted in Figure 2.

CropScap data are obtained by classifying high-resolution aerial images collected throughout the year across the USA (USDA, 2017). Therefore, the published data are not error free. The classification accuracy (user's accuracy and producer's accuracy) is presented in Table 1 (USDA, 2017). Except for Alfalfa, a class with a producer's accuracy of 90%, all classes had to be carefully investigated before using them for this study.

To improve the quality of this sample set, we removed the wrongly classified samples by implementing the methodology presented by Belgiu, Csillik (2018) and generated randomly 100 samples per class. We computed their phenological profiles, i.e. growth cycle, assessed the differences between these resulting profiles and the crop calendar and discarded all samples that deviate from the pattern of the represented crops. In the end, we had 50 training samples per class.

Class	Producer's accuracy	User's accuracy
Alfalfa	91.60%	87.50%
Fallow	80.80%	80.50%
Other Hay	53.60%	65.50%
Onion	78.60%	71.40%
Sugar beet	46.70%	86.00%
Winter Wheat	68.10%	72.40%
Lettuce	20.10%	45.10%

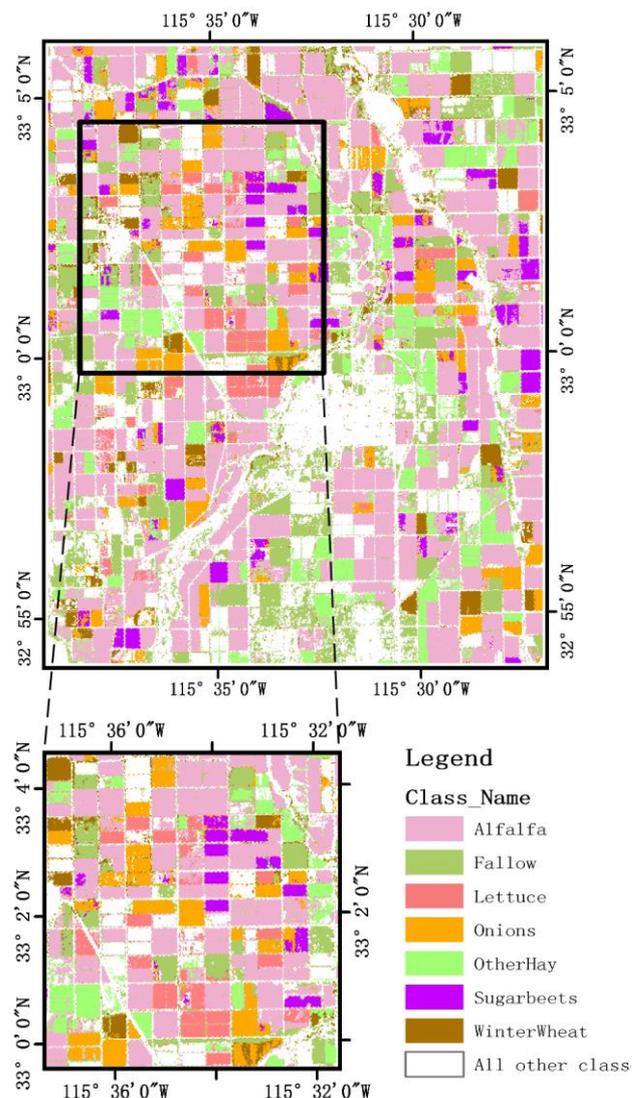Table 1. Accuracy of CropScape - Cropland Data Layer (CDL)

Figure 2. Spatial distribution of target crops in the study area. The legend was adopted from the USDA (USDA, 2017)

The validation samples were generated using the same procedure as explained above. Careful attention was paid to the spatial distribution of these data and exclude the agricultural plots used to generate the training samples.

3.4 Classification accuracy

Given the limitation of the widely used kappa index for remote sensing applications as reported by Foody (2020), we used overall accuracy together with user's accuracy and producer's accuracy metrics to evaluate the quality of the classification results obtained by using the three methods (Congalton, 1991).

4. RESULTS AND DISCUSSION

TWDTW outperformed both DTW and WDDTW, obtaining an overall accuracy of 88% as compared to 67% and 57% accuracy obtained by WDDTW and DTW respectively.

TWDTW performed better than the other two evaluated methods because of its capability to reduce considerably the misclassification rate of the crops with a large time lag on the temporal NDVI sequences. For example, the temporal patterns of the investigated crops (Figure 1) show that winter wheat and onions have similar duration and amplitude of their NDVI values. The only difference consists in the fact that the peak of vegetation of winter wheat occurs one month earlier than those of onions. DTW and WDDTW do not account for this difference and, therefore, the classification errors for these two crops are high. TWDTW, on the other hand, obtained a higher user's accuracy for winter wheat (82%) as compared to DTW and WDDTW which obtained an user's accuracy of only 30% and 40%, respectively.
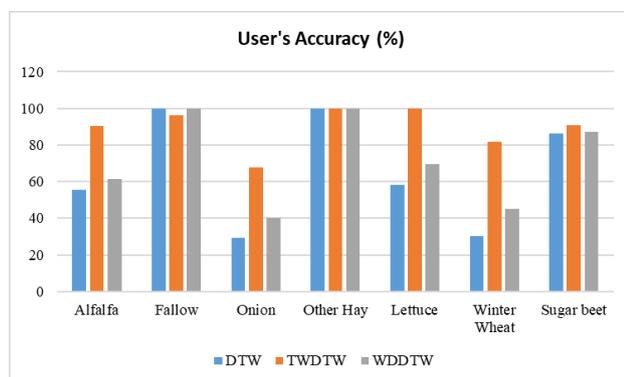

Figure 3. User's accuracy obtained by applying DTW-Dynamic Time Warping, TWDTW- Time Weighted Dynamic Time Warping, WDDTW- Weighted Derivative DTW

TWDTW achieved good classification results as quantified by both user's and producer's accuracy for the following classes: alfalfa, fallow, other hay, and sugar beets (Figure 3 and Figure 4). The best classified classes are fallow (producer's accuracy of 100% and user's accuracy of 96%) and alfalfa with a producer's accuracy of 94% and an user's accuracy of 90%. Onions yielded a producer's and user's accuracy of 76% and 70% respectively. WDDTW outperformed DTW by 10% (67% vs. 57%). It performs better than DTW for the classification of sugar beets (producer's accuracy of 80%) or double cropping such as lettuce (Figure 2 and Figure 3), obtaining a producer's accuracy of 64% and an user's accuracy of 70%.

DTW is computationally very demanding. To overcome this limitation, the input satellite images time series should be reduced to those images that are relevant for the investigated crops. Previous studies proved that identifying the optimal temporal window for crops mapping does not only reduce the computational time, but it also improves the classification results (Meng et al., 2020). The optimal temporal window can be selected using either a data-driven or a knowledge-based approach, i.e. using the crop calendar of the cultivated crops in the investigated agricultural regions.
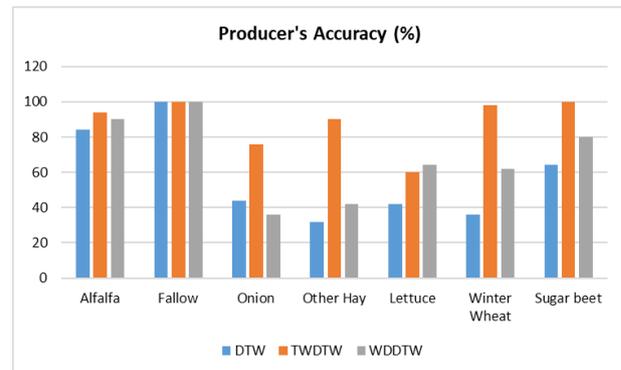

Figure 4. Producer's accuracy obtained by applying DTW-Dynamic Time Warping, TWDTW- Time Weighted Dynamic Time Warping, WDDTW- Weighted Derivative DTW

The quality of the DTW-based classification results depends on the quality of the training samples. Collecting training samples through fields campaign for a large number of cops is a time-consuming and expensive task (Maxwell et al. 2018). Therefore, we need efficient solutions to generate training samples to support the production of timely and reliable crop information. Transfer learning (Tuia et al., 2011), crowdsourcing initiatives (Fritz et al. 2009), and utilization of existing inventories to guide the labeling of the new training samples (Huang et al. 2020) are promising solutions to address this challenge.

5. CONCLUSIONS

This study evaluated the performance of DTW and two of its variations, TWDTW and WDDTW for crops mapping from Sentinel-2 time series mapping. TWDTW proved to be the most efficient among them due to the additional time constraint that can be defined when matching two temporal sequences representing the crops growth cycles. In future work, we plan to investigate a knowledge-based selection of the optimal temporal window for each crop when applying WDDTW.

REFERENCES

Belgiu, M., Csillik, O., 2018. Sentinel-2 cropland mapping using pixel-based and object-based time-weighted dynamic time warping analysis. *Remote Sensing of Environment,* 204, 509-523.

Boryan, C., Yang, Z., Mueller, R., Craig, M., 2011. Monitoring US agriculture: the US Department of Agriculture, National Agricultural Statistics Service, Cropland Data Layer Program. *Geocarto International,* 26(5), 341-358.

Brelsford, C., Martin, T., Hand, J., Bettencourt, L.M., 2018. Toward cities without slums: Topology and the spatial evolution of neighborhoods. *Science advances,* 4(8), eaar4644.
Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment,* 37(1), 35-46.

Csillik, O., Belgiu, M., Asner, G.P., Kelly, M., 2019. Object-Based Time-Constrained Dynamic Time Warping Classification of Crops Using Sentinel-2. *Remote Sensing,* 11(10), 1257.

Dahal, D., Wylie, B., and Howard, D. 2018. Rapid Crop Cover Mapping for the Conterminous United States. *Scientific Reports*, 8, 8631.

Foody, G.M., 2020. Explaining the unsuitability of the kappa coefficient in the assessment and comparison of the accuracy of thematic maps obtained by image classification. *Remote Sensing of Environment,* 239, 111630.

Fritz, S., See, L., McCallum, I., You, L., Bun, A., Moltchanova, E., Duerauer, M., Albrecht, F., Schill, C., Perger, C., 2015. Mapping global cropland and field size. *Global change biology,* 21(5), 1980-1992.

Fritz, S., McCallum, I., Schill, C., Perger, C., Grillmayer, R., Achard, F., Kraxner, F., and Obersteiner, M. 2009. Geo-Wiki. Org: The use of crowdsourcing to improve global land cover. *Remote Sensing*, 1(3), 345-354.

Guan, X., Huang, C., Liu, G., Meng, X., Liu, Q., 2016. Mapping rice cropping systems in Vietnam using an NDVI-based time-series similarity measurement based on DTW distance. *Remote Sensing,* 8(1), 19-23.

Guan, X., Liu, G., Huang, C., Meng, X., Liu, Q., Wu, C., Ablat, X., Chen, Z., Wang, Q.,2018. An Open-Boundary Locally Weighted Dynamic Time Warping Method for Cropland Mapping. *ISPRS International Journal of Geo-Information,* 7(2), 75-87.

Huang, H., Wang, J., Liu, C., Liang, L., Li, C., and Gong, P. 2020. The migration of training samples towards dynamic global land cover mapping. *ISPRS Journal of Photogrammetry and Remote Sensing*, 161, 27-36

Jeong, Y.-S., Jeong, M.K., Omitaomu, O.A., 2011. Weighted dynamic time warping for time series classification. *Pattern Recognition,* 44(9), 2231-2240

Johnen, B., Kuhlenkoetter, B., 2016. A Dynamic Time Warping algorithm for industrial robot motion analysis. In: *2016 Annual Conference on Information Science and Systems (CISS)*, pp. 18-23.

Keogh, E.J., Pazzani, M.J., 2001. Derivative dynamic time warping. *Proceedings of the 2001 SIAM international conference on data mining*: *SIAM*. Society for Industrial and Applied Mathematics.

Maus, V., G, C., Cartaxo, R., Sanchez, A., Ramos, F.M., Queiroz, G.R.d., 2016. A Time-Weighted Dynamic Time Warping Method for Land-Use and Land-Cover Mapping. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing,* 1(99), 1-11.

Maxwell, A.E., Warner, T.A., and Fang, F. 2018. Implementation of machine-learning classification in remote sensing: an applied review. *International Journal of Remote Sensing*, 39, 2784-2817.

Meier, J., Zabel, F., and Mauser, W. 2018. A global approach to estimate irrigated areas–a comparison between different data and statistics. *Hydrology and Earth System Sciences*, 22, 1119-1133.

Meng, S., Zhong, Y., Luo, C., Hu, X., Wang, X., and Huang, S. 2020. Optimal Temporal Window Selection for Winter Wheat and Rapeseed Mapping with Sentinel-2 Images: A Case Study of Zhongxiang in China. *Remote Sensing*, 12(2), 226.

Petitjean, F., Inglada, J., Gançarski, P., 2012. Satellite image time series analysis under time warping. *IEEE Transactions on geoscience and remote sensing,* 50(8), 3081-3095.

Sakoe, H., Chiba, S., 1978. Dynamic programming algorithm optimization for spoken word recognition. *IEEE transactions on acoustics, speech, and signal processing,* 26(1), 43-49.

Salmon, J.M., Friedl, M.A., Frolking, S., Wisser, D., and Douglas, E.M. 2015. Global rain-fed, irrigated, and paddy croplands: A new high resolution map derived from remote sensing, crop inventories and climate data. *International Journal of Applied Earth Observation and Geoinformation*, 38, 321-334

Tsevas, S., Iakovidis, D.K., 2010. Dynamic time warping fusion for the retrieval of similar patient cases represented by multimodal time-series medical data. *Proceedings of the 10th IEEE International Conference on Information Technology and Applications in Biomedicine*. doi: 10.1109/ITAB.2010.5687649.

Tuia, D., Volpi, M., Copa, L., Kanevski, M., and Munoz-Mari, J. 2011. A survey of active learning algorithms for supervised remote sensing image classification. *IEEE Journal of Selected Topics in Signal Processing*, 5(3), 606-617.

United-Nations, 2015. World population prospects: the 2015 revision. *Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat.* New York: Department of Economic and Social Affairs.

USDA, 2017. California Agricultural Statistics. Annual Bulletin. Available online at http://www.nass.usda.gov/Statistics_by_State/California/Public ations/California_Ag_Statistics/2013cas-all.pdf (1 May 2020).

Waldner, F., Canto, G.S., and Defourny, P. 2015. Automated annual cropland mapping using knowledge-based temporal features. *ISPRS Journal of Photogrammetry and Remote Sensing*, 110, 1-13.

Wang, S., So, E., Smith, P., 2015. Detecting tents to estimate the displaced populations for post-disaster relief using high resolution satellite imagery. *International Journal of Applied Earth Observation and Geoinformation,* 36(1), 87-93.

Xiong, J., Thenkabail, P.S., Gumma, M.K., Teluguntla, P., Poehnelt, J., Congalton, R.G., Yadav, K., and Thau, D. 2017. Automated cropland mapping of continental Africa using Google Earth Engine cloud computing. *ISPRS Journal of Photogrammetry and Remote Sensing,* 126, 225-244

Yan, L., Roy, D.P., 2014. Automated crop field extraction from multi-temporal Web Enabled Landsat Data. Remote Sensing of Environment, 144, 42-64.

Zhong, L., Hu, L., Zhou, H., Tao, X., 2019. Deep learning based winter wheat mapping using statistical data as ground references in Kansas and northern Texas, US. *Remote Sensing of Environment,* 233, 111411.