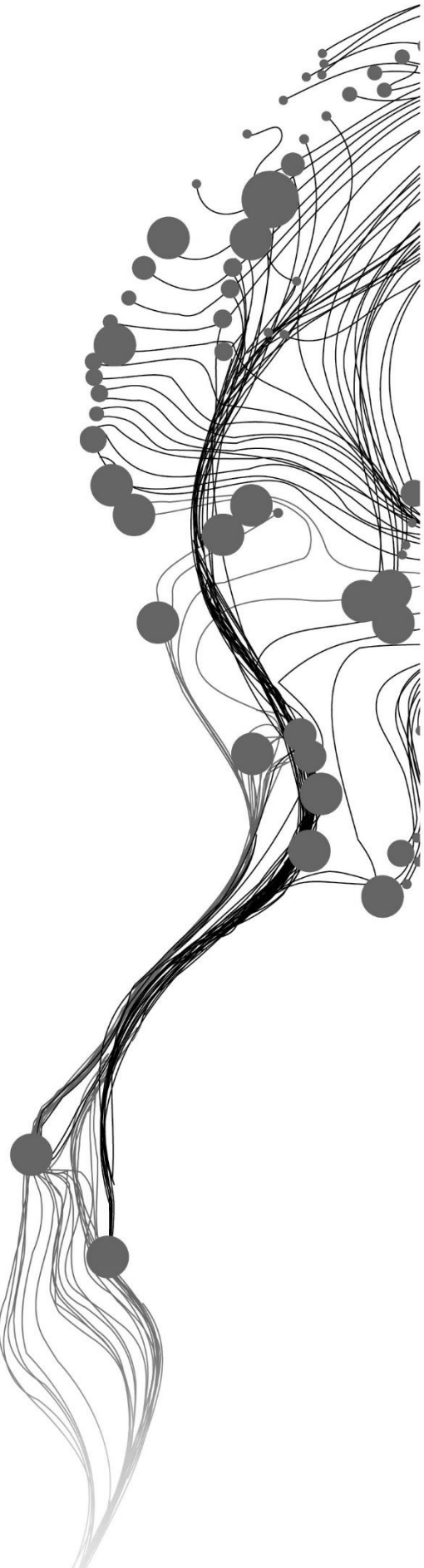


Reduction of Computation Time of Dynamic Time Warping Based Methods used for Cropland Mapping

YIBO ZHOU
February, 2019

SUPERVISORS:
Dr. M. Belgiu
Dr. M.T. Marshall



Reduction of Computation Time of Dynamic Time Warping Based Methods used for Cropland Mapping

YIBO ZHOU

Enschede, The Netherlands, February, 2019

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Geoinformatics

SUPERVISORS:

Dr. M. Belgiu

Dr. M.T. Marshall

THESIS ASSESSMENT BOARD:

Prof. dr. ir. A. Stein (Chair)

Dr. V. Maus (External Examiner, Institute for Ecological Economics,
Vienna University of Economics and Business, Austria)

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

The efficiency of cropland mapping is an essential criterion for the implementation of sustainable agricultural practices and crop monitoring. With the increasing spatial and temporal resolution of satellite images, efficient cropland mapping calls for the satellite image time series (SITS) classification methods that are both accurate and fast. Dynamic time warping (DTW) and its variations showed significant advantages in coping with the irregular time series data. This MSc thesis evaluates the performances of DTW, time-weighted DTW (TWDTW), weighted derivative DTW (WDDTW) and non-isometric transform DTW (NTDTW) in terms of accuracy and computation time for time series classification. The TWDTW achieved the highest overall accuracy which was 88.3%, and was the most time-consuming method. Warping path window was used to accelerate the computation of DTW-based methods. The TWDTW was selected to define the warping path window for reducing the computation time. Without the overall accuracy changing, the computation time of TWDTW reduced by 56.5%. Another method for reducing computation time called pattern decomposition was applied on both the TWDTW and TWDTW with warping path window. With a slightly changed overall accuracies, the computation time reduced 67.7% and 75.4% respectively.

Keywords

Remote sensing, Agriculture, Satellite image time series, Dynamic time warping, Warping path window

ACKNOWLEDGEMENTS

Foremost, I thank my parents and my relatives for supporting my MSc in the Netherlands not only on the financial but also the emotional encouragement.

I would like to say a big thank you to my first supervisor dr. M. Belgiu (Mariana) for her invaluable support and guidance throughout my thesis. My gratitude also goes to my second supervisor, dr. M.T. Marshall (Michael), for his patience to give me ideas, advices and support me in the academics writing aspect. This thesis would not have been possible without their scientific knowledge and constructive advices. There was so much fun and freedom working with the Mariana and Michael.

I thank all my Geoinformatics classmates. You built a wonderful memory since the modules until thesis here, especially thanks for Ratna Mayasari, Mwanaidi Mohamed, Khairya Masoud, Yan Zhang, Jianda Yan for helping me a lot during the study period.

Finally, sincere thanks to all of my friends in the Netherlands. I cannot have a colorful life here without you. Thanks especially to Maryam, Gerardo, Wei Li for giving me emotional encouragement. I am really grateful that Maomao has become the reason for me to study for MSc degree in the Netherlands.

TABLE OF CONTENTS

1.	Introduction.....	1
1.1.	Motivation and Problem Statement.....	1
1.2.	Research Objectives.....	2
1.3.	Research Questions.....	2
2.	The State of the art.....	4
2.1.	Variations of DTW.....	4
2.2.	Time-Weighted Dynamic Time Warping (TWDTW).....	4
2.3.	Weighted Derivative Dynamic Time Warping (WDDTW).....	5
2.4.	Non-isometric Dynamic Time Warping (NTDTW).....	6
2.5.	DTW warping path window.....	7
2.5.1.	Sakoe-Chiba Band.....	8
2.5.2.	Itakura Parallelogram.....	8
3.	Materials and methods.....	10
3.1.	Study area.....	10
3.2.	Data collection and pre-processing.....	11
3.3.	Methods.....	12
3.3.1.	Overview.....	12
3.3.2.	Training and validation samples.....	13
3.3.3.	Temporal phenological patterns.....	14
3.3.4.	Classification.....	15
3.3.5.	DTW.....	16
3.3.6.	TWDTW.....	17
3.3.7.	WDDTW.....	18
3.3.8.	NTDTW.....	19
3.3.9.	Defining the warping path window width.....	19
3.3.10.	Decomposition of phenological patterns.....	20
4.	Results and discussion.....	21
4.1.	Results.....	21
4.1.1.	Phenological patterns decision.....	21
4.1.2.	Classification experiments on three DTW-based methods.....	23
4.1.3.	Defining the warping path window.....	25
4.1.4.	Decomposition of phenological patterns.....	26
4.2.	Discussion.....	29
4.2.1.	Selection of the time series interval.....	29
4.2.2.	Assessment of three DTW-based methods.....	29
4.2.3.	Defining the warping path window.....	32
4.2.4.	Decomposition of phenological patterns.....	32
4.3.	Conclusions.....	33
	References.....	35

LIST OF FIGURES

Figure 1 Local differences between two sequences (Keogh & Pazzani, 2001).....	6
Figure 2 Sakoe-Chiba Band and Itakura Parallelogram warping window (Ratanamahatana & Keogh, 2004).....	8
Figure 3 Study area in the Imperial Valley of California, USA	10
Figure 4 Methods flowchart	13
Figure 5 Study area.....	16
Figure 6 DTW distance matrix and Warping path window (Jeong et al., 2011).....	17
Figure 7 MLWF with different α	18
Figure 8 Temporal profiles of different dates selection of Sentinel-2 time series data.....	22
Figure 9 Classification results of DTW with warping path window width from 5 to 1 and TWDTW with warping path window.	26
Figure 10 Classification results of pattern decomposition DTW (PD-DTW), pattern decomposition DTW with warping path window (PD-DTW-W), pattern decomposition TWDTW (PD-TWDTW), pattern decomposition TWDTW with warping path window (PD-TWDTW-W). ..	27
Figure 11 Comparison of original patterns and decomposed patterns.....	28

LIST OF TABLES

Table 1: Spectral bands for the Sentinel-2 sensors (ESA, 2018)	11
Table 2: Dates list of the Sentinel-2 time series data used in this research.....	12
Table 3 List of the first 9 crops sorted descending by pixels count.....	14
Table 4 Overall accuracies and computation time of three periods' DTW classifications.....	22
Table 5 Classification accuracy of DTW method from September 2016 to September 2017.....	23
Table 6 Classification accuracy assessments of TWDTW, WDDTW and NDTW methods.....	24
Table 7 The Overall Accuracy and computation time of DTW and TWDTW methods with different warping path window widths.	25
Table 8 Accuracy results and computation time of DTW and TWDTW with and without patterns decomposition classification.....	28

1. INTRODUCTION

1.1. Motivation and Problem Statement

With the rapidly increasing population of the world, agriculture, as the food supply, is facing enormous challenges in the aspects of production and management (Waldner et al., 2015). Furthermore, the increasing frequency and intensity of extreme events such as drought, flooding, and high maximum temperatures are hindering agricultural development (Edenhofer et al., 2015). The development of sustainable natural resources management calls for the improvement of cropland mapping and monitoring methodologies (Matton et al., 2015). Offering timeliness, global coverage and objective observation, satellite remote sensing has become a critical source of data for cropland mapping and monitoring (Atzberger & Clement, 2013). Satellite image time series (SITS) data are demonstrated as being useful for describing trends or discrete change events of environmental phenomena (Gómez et al., 2016).

According to Petitjean et al. (2012), time series cropland mapping has three challenges: (1) the lack of samples used to train the supervised algorithm; (2) missing temporal data caused by clouds obscuration; (3) annual changes of phenological cycles caused by weather or by variations in the agricultural practices. Dynamic Time Warping (DTW) (Sakoe, 1978) has shown well performance in coping with these challenges (Baumann et al., 2017; Petitjean et al., 2012). The DTW method works well for shape matching but is not suited for time series classification. It disregards the temporal range when applying the similarity measure between two time series (Maus et al., 2016). Each crop class has a distinct phenological temporal profile although there are internal variations (Reed et al., 1994). However, the phenological cycles in practice will not vary extremely. The temporal range and amplitude of the cycle vary in a range. In the DTW method, the similar amplitude cycles will be matched without considering their temporal durations or phase differences. Consequently, there will be two kinds of wrongly matchings: (1) two phenological cycles which have similar amplitude, but with different durations. These two cycles are supposed to be different crop types while the DTW method will match them together. (2) two phenological cycles which have a similar shape but start in different seasons during the year. The DTW method will match these two crops. In our study, we use the DTW and other DTW-based methods to classify the temporal profiles of seven crops which contained varying degree of internal variations. The DTW and DTW-based methods showed the abilities to endure this kind of variations in temporal profiles (Bagnall et al., 2017; Górecki & Łuczak, 2014; Jeong et al., 2011).

Although the DTW and its variations have shown their advantages in the classification accuracy of time series data, all DTW-based methods are facing a problem that the computation time is long. In Belgiu & Csillik (2018)'s experiments, they applied Random Forest and a variation of DTW called Time-Weighted

Dynamic Time Warping (TWDTW) (Maus et al., 2016) on three different study areas. In the computation time statistics, TWDTW required much longer time than Random Forest method. For example, for their study area 3, TWDTW spent 30 hours for classification, while the Random Forest only required 1 hour for 25 iterations (Belgiu & Csillik, 2018). The long computation time of DTW method is reasonable due to the uncertainty of the warping extent in each two sequences' DTW distance calculation (Gullo et al., 2009). The original DTW method calculates the DTW distance between all the points between these two sequences, which takes long computation time. There were many studies tried to improve the DTW method's efficiency in different aspects. Petitjean et al. (2015) used the Nearest Centroid algorithm to average the set of sequences first, and then applied the DTW method for classification. This kind of reducing computation time method reduces the quantity of input time series data so as to reduce the computation time. With the same idea, Belgiu & Csillik (2018) adopted an object-based approach on the input satellite image time series data. The input images were first segmented and then applied for the subsequent classification step. In this way, the input data were greatly simplified. On the other hand, there were also many studies concentrated on the DTW algorithm's simplification. Sakoe (1978) noticed that in the calculation of DTW distance between two sequences, there are computing redundancy. The definition of warping path window could greatly reduce this kind of computing redundancy (Sakoe, 1978). In our study, we applied the warping path window as one of our methods to reduce the computation time of DTW and DTW-based methods.

1.2. Research Objectives

The overall objective of this research is to reduce the computation time of DTW and DTW-based methods for cropland mapping from satellite image time series.

The Specific objectives are:

- To assess the performance of DTW and three DTW-based methods for cropland mapping.
- To test different widths of the warping path window defined on the DTW-based methods.
- To reduce the computation area by decomposing the patterns refer to the phenology of target crops.

1.3. Research Questions

- To what extent do these three DTW-based methods improve the classification accuracy of original DTW method in the study area?
- What is the computational time required by the DTW and the three DTW-based methods?
- What is the computational time required by the different widths for defining the warping path window on the DTW and variations?
- How the widths of the warping path window affect the classification accuracies of DTW and DTW-based methods?
- How to decompose the phenological pattern of each crop by considering the phenology in a mathematical formula or a logical method?

- How the decomposition of phenological patterns affects the accuracy and computation time of DTW and DTW-based method?

2. THE STATE OF THE ART

The state of the art is divided into two parts. Part 1 presents the main variations of Dynamic Time Warping (DTW) similarity measurement used in Satellite Image Time Series (SITS). These DTW-based variations achieved better accuracy of SITS classification. Part 2 presents the two most frequently used methods of defining warping path window. The warping path window not only alleviates wrongly warpings, but it also speeds up the DTW distance calculation.

2.1. Variations of DTW

DTW is an algorithm that calculates an optimal match between two sequences for measuring similarity. DTW differs from Euclidean distance (ED) by allowing one point on one sequence to be matched to many points on the other sequence and by allowing time lag between the matched points. The DTW was first proposed in the speech recognition field (Sakoe, 1978) and was introduced into the time series comparison to find patterns (Berndt & Clifford, 1994). Petitjean et al. (2012) first demonstrated that the DTW has advantages in coping with three challenges in SITS classification: (1) the lack of samples used to train the supervised algorithm; (2) missing temporal data caused by clouds obscuration; (3) annual changes of phenological cycles caused by weather or by variations in the agricultural practices. There are three variations of DTW method that have their advantages in coping with these challenges. These methods are presented below.

2.2. Time-Weighted Dynamic Time Warping (TWDTW)

There is a variation of the DTW method called the Time-Weighted Dynamic Time Warping (TWDTW) (Maus et al., 2016), which adopts a weight into the calculation of DTW distance. The weight is controlled by a Modified Logistic Weight Function (MLWF). The MLWF takes two matching points' phase difference into consideration. 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. The weight can be tuned to suit different situations encountered 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. This method can be successfully used for agriculture mapping where different agriculture practices might cause variations in the growth cycles of the same crop. For example, due to different cultivation practices, the same crop might be planted or harvested in 15-30 days' time difference (USDA, 2010). In this situation, the user can define a time difference tolerance between crop cycles in two sequences to suit different study area's cultivation practices.

The TWDTW can alleviate two kinds of errors that might occur when applying the DTW method: (1) two phenological cycles with similar amplitude, but with different durations. Despite the fact that these two cycles characterize two different crop types, the DTW method will match them together. (2) two

phenological cycles with a similar shape, but which start in different seasons during the year. These two cycles are supposed to be the same crop which is cultivated in different seasons (e.g. spring wheat and winter wheat). The DTW method will match these two crops. Thus, the time series cropland classification method needs to balance between shape matching and temporal alignment. The weight in TWDTW is a time constraint which controls the time warping extent and makes the time series alignment dependent on the seasons.

2.3. Weighted Derivative Dynamic Time Warping (WDDTW)

Another variation of the DTW method is called Weighted Derivative Dynamic Time Warping (WDDTW). This method combines the idea of TWDTW and Derivative Dynamic Time Warping (DDTW) (Keogh & Pazzani, 2001). The DDTW does not use the value of the sequences directly but uses the first derivative of the sequences to calculate their DTW distance. This idea is adopted to WDDTW. The WDDTW defines a weight to calculate the distance between two sequences. However, the distance between two sequences is not calculated directly on the original value. A neighbour point averaging is applied to the original sequences. In this way, two new sequences will be created. The averaging is simply the average of the slope of the line through the test point and its left neighbour, and the slope of the line through the left neighbour and the right neighbour. In the next step, the two new sequences will be matched following the procedure implemented in TWDTW and explained above.

This WDDTW method tries to solve one of the weaknesses of DTW, namely that the DTW only considers the difference of Y-axis value of the points from two sequences. For example, assuming that two points A and B have identical values, but A is part of a rising trend and B is part of a falling trend. DTW considers these two points as being similar and matches them together, although A and B are not supposed to be matched. To prevent this problem, the first derivative of the sequences is considered. There is also a problem of the TWDTW method that can be alleviated by the WDDTW method. For example, the two sequences depicted in Figure 1a are similar. One of the sequences is modified to a deeper valley as depicted in Figure 1c. Therefore, the DTW distance between the two sequences in this particular location will be higher. If these two sequences have also a time lag, the TWDTW distance will be higher than DTW distance due to the weight. However, from the results of the distance alone, we cannot distinguish whether the higher distance is caused by the shape difference or by the time lag. The WDDTW uses the first derivative to acquire the shape features of the two sequences and to get a better matching between them.

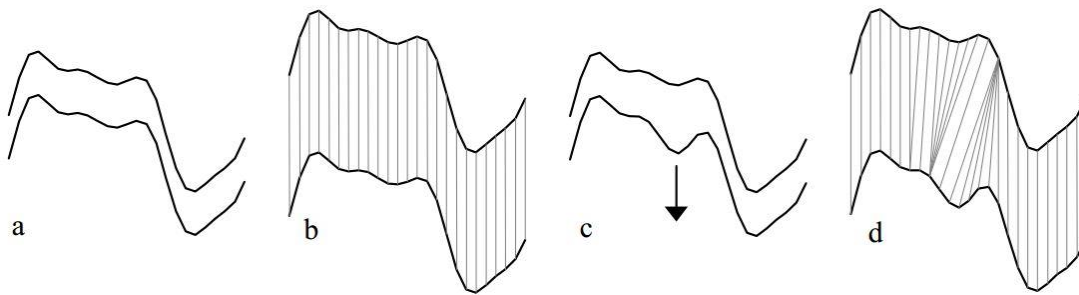


Figure 1 Local differences between two sequences (Keogh & Pazzani, 2001)

2.4. Non-isometric Dynamic Time Warping (NTDTW)

Besides the two above-mentioned variations of DTW, there is also the so-called Non-Isometric Transform Dynamic Time Warping (NTDTW) (Jeong et al., 2011). This method uses the sequences' mathematical transforms to calculate the DTW distance and combines the transform DTW distance with original DTW distance by using different weights. The mathematical transforms are commonly used in the classification of time series to extract higher order features (Faloutsos et al., 1994). The transforms are Cosine, Sine, and Hilbert. The transforms convert the original sequences from the time domain to the frequency domain, which contains more shape features of the original sequences. The weight controls the effect of the transformed DTW distance. Similar to the WDDTW method that takes the shape features of the sequences into account, the NTDTW method tries to improve the matching accuracy between sequences where cycles have internal variation. The WDDTW and NTDTW all claimed that they perform better than the original DTW method by considering the sequences' higher order features (Górecki & Łuczak, 2014; Jeong et al., 2011). However, there is no directly performance comparison between WDDTW and NTDTW yet.

In this study, there will be a performance comparison among TWDTW, WDDTW and NTDTW for cropland mapping. The TWDTW, WDDTW, and NTDTW methods have been successfully implemented and produced high classification accuracy (Górecki & Łuczak, 2014; Jeong et al., 2011; Maus et al., 2016). The TWDTW has been applied to a cropland classification in a tropical forest area with an overall accuracy of 87.32%, which is higher than DTW method's 70.14% (Maus et al., 2016). In Jeong et al.'s, (2011) study, they compared WDDTW and DTW using different time series datasets. Most of the error rates of WDDTW is lower than the error reported for the DTW (Jeong et al., 2011). The NTDTW method was experimented on 47 time series datasets (E. Keogh et al., 2011). By comparing with the conventional DTW method, there is a significant reduction in error rate for most datasets (Górecki & Łuczak, 2014). All of these three methods can be adapted to different crop types. The tuneable parameters in these methods help them adapt to specific variations of phenological cycle shapes. This is meaningful to SITS cropland mapping.

DTW and DTW-based methods have been applied to cropland mapping after the advantages (mentioned at the beginning of Part I) of DTW and DTW-based methods in dealing with SITS land-use classification were demonstrated (François Petitjean et al., 2012). Petitjean et al. (2012) applied the DTW method to Formosat-2 images classification in order to deal with issues raised by high resolution SITS, namely the

irregular sampling caused by clouds obscuration, ground truth missing, weather and variations of agricultural practices. Guan et al. (2016) applied the DTW to NDVI time-series from MODIS for mapping rice cropping system and confirmed that DTW is suitable for areas in which rice planting scheduling is flexible and rice growth periods in different phases are uncertain. Maus et al. (2016) applied the TWDTW to classify crops with various vegetation dynamics by EVI derived from MODIS data. The TWDTW showed well performance in identifying single cropping, double cropping, forest and pasture. Belgiu & Csillik (2018) applied TWDTW to NDVI time series derived from Sentinel-2 data cropland mapping on three study areas. The TWDTW method showed well performance in the crop type classifications. The highest overall accuracy among the classifications reached 96.19% (Belgiu & Csillik, 2018). Guan et al. (2018) applied an open-boundary locally weighted dynamic time warping (OLWDTW) distance method to NDVI time-series from MODIS for cropland mapping in Southeast Asia. There was a comparison between the OLWDTW and DTW. The accuracy of OLWDTW was more than 5% higher than DTW's (Guan et al., 2018).

2.5. DTW warping path window

The classic DTW method is time-consuming (Belgiu & Csillik, 2018; Petitjean et al., 2012). The three DTW-based methods described above rely on complicated internal calculations which makes them even more time-consuming. The TWDTW method has two parameters i.e. half-length of the sequence and the level of penalization for the points with larger phase difference, to be tuned to suit different growth cycles of the investigated crop types. The WDDTW method has a parameter for weighting the derivative distance to be tuned. The NTDTW has also a parameter for weighting transform distance to be tuned. The optimal parameters are defined by assessing the classification accuracy achieved by defining different parameter values. The tuning of these parameters increases the computational time. Therefore, there is an obvious need to reduce the computational time of these methods.

Several research efforts are dedicated to defining the width of DTW warping path window (Itakura, 1975; Sakoe, 1978; Ratanamahatana & Keogh, 2004; Sakurai et al., 2007; Silva & Batista, 2016; Tan et al., 2018). The defined width constraints the DTW distance computation. Therefore, only the DTW distance computations within the width of the defined window will be taken. The width constraint helps reduce the computation cost by narrowing the search for a qualified path. In addition to reducing the DTW distance computation time, the warping path window also alleviates wrongly warping situations such as those when a point of one sequence matches the point which has a bigtime lag on another sequence.

The definitions of warping path window can be divided into two kinds: (1) global constraints and (2) local constraints. There are two most frequently used methods for defining warping path window, namely the Sakoe-Chiba band (Sakoe, 1978) and the Itakura Parallelogram (Itakura, 1975). They both give the calculation of DTW distance global constraints. The global constraint gives the DTW distance calculation a whole limit on how far the warping path may stray from the diagonal of the warping matrix (Das, 1978; Guan et al., 2016; Myers, Rabiner, & Rosenberg, 1980).

There are also some local constraints methods to define the warping path window (Anh et al., 2018; Ratanamahatana & Keogh, 2004; Silva & Batista, 2016; Tan et al., 2018). These methods try to define the warping path window by extracting each element's warping extent from diagonal along the warping path of the two sequences' DTW distance calculation. The constraints will vary with the calculation of each DTW distance, which is more complicated.

2.5.1. Sakoe-Chiba Band

The Sakoe-Chiba Band was proposed for speech patterns aligning (Sakoe, 1978). The warping window gives the DTW distance matrix global constraints by defining the window width. The width stands for the maximum warping distance of a point on one sequence that can be matched to the other sequence. The similarity between two sequences will be assessed only within the defined warping window as depicted in Figure 2. In Sakoe-Chiba Band approach, the warping path window has the same width along the diagonal of the DTW distance matrix. It means every points of the sequence are allowed warping to the points on the other sequence within the defined width. The width is a constant. The Sakoe-Chiba Band is widely used not only in speech patterns aligning but also in music (Ning Hu, 2003; Zhu, Shasha, & Zhao, 2003), finance (Berndt & Clifford, 1994), medicine (Gollmer & Posten, 1995), biometrics (Deeb, Jannetta, Rosenbaum, Kerber, & Drayer, 1979) and robotics (Deeb et al., 1979).

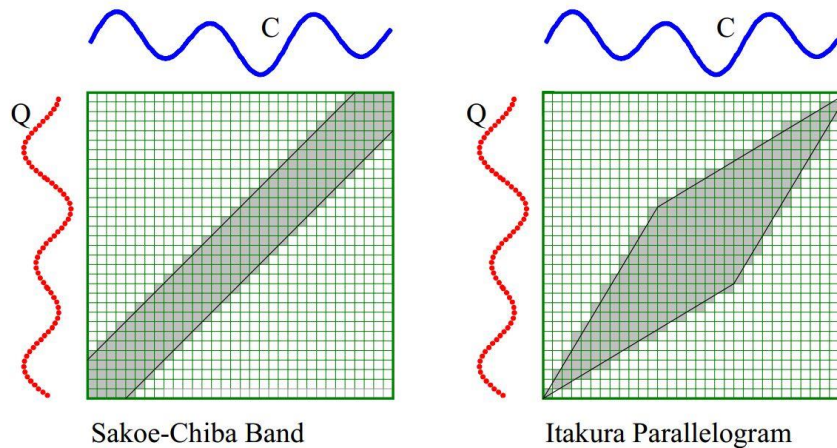


Figure 2 Sakoe-Chiba Band and Itakura Parallelogram warping window (Ratanamahatana & Keogh, 2004)

2.5.2. Itakura Parallelogram

The Itakura Parallelogram was also proposed for speech recognition (Itakura, 1975). In this method, the warping path window is more flexible to address the needs of different warping extents. In some cases, the warpings are not significant at the beginning and the end of the assessed time sequences. The parallelogram window's shape can be tuned with two parameters controlling the maximum warping width and the slope. The slope stands for the warping extent's increasing speed from the beginning of the warping path to the maximum warping width. The setting of the window parameters might be designed based on domain knowledge or experience.

Compare to Sakoe-Chiba Band, the Itakura Parallelogram is more fault-tolerant. The rigorous window width in Sakoe-Chiba Band may cause the optimal warping path to be cut by the window boundary which in turns might affect the classification accuracy. Nevertheless, the Sakoe-Chiba Band is simpler to apply in the calculation than the Itakura Parallelogram. The tuning of the parameters in Itakura Parallelogram is more time consuming.

3. MATERIALS AND METHODS

3.1. Study area

The study area of this research is in Imperial County, southern California, United States, from 33.1°N, 115.65°W to 32.9°N, 115.45°W. Figure 3(a) illustrates the location of the study area. Imperial County is one of the highest-producing counties for sugar beets (100%), onions (22%), hay (alfalfa, over 17%), lettuce (over 17%) and wheat (over 15%) (USDA, 2017a). Figure 3(b) shows the false color combination (band 8 near-infrared, band 4 red, band 3 green) for California in May 11, 2017. All the images in this study were projected in WGS 1984 UTM Zone 11N.

As the Mediterranean climate in this study area, the rain falls mostly outside the crops growing season. The agriculture is heavily dependent on irrigated water from the Colorado River. For the year 2017, the average precipitation was 61.72 mm, and the annual mean temperature was above 27°C (USDA, 2017a). Seven classes were selected for analysis, each of them occupying at least 3% of the study area: alfalfa, fallow, other hay, onions, sugar beets, winter wheat and lettuce.

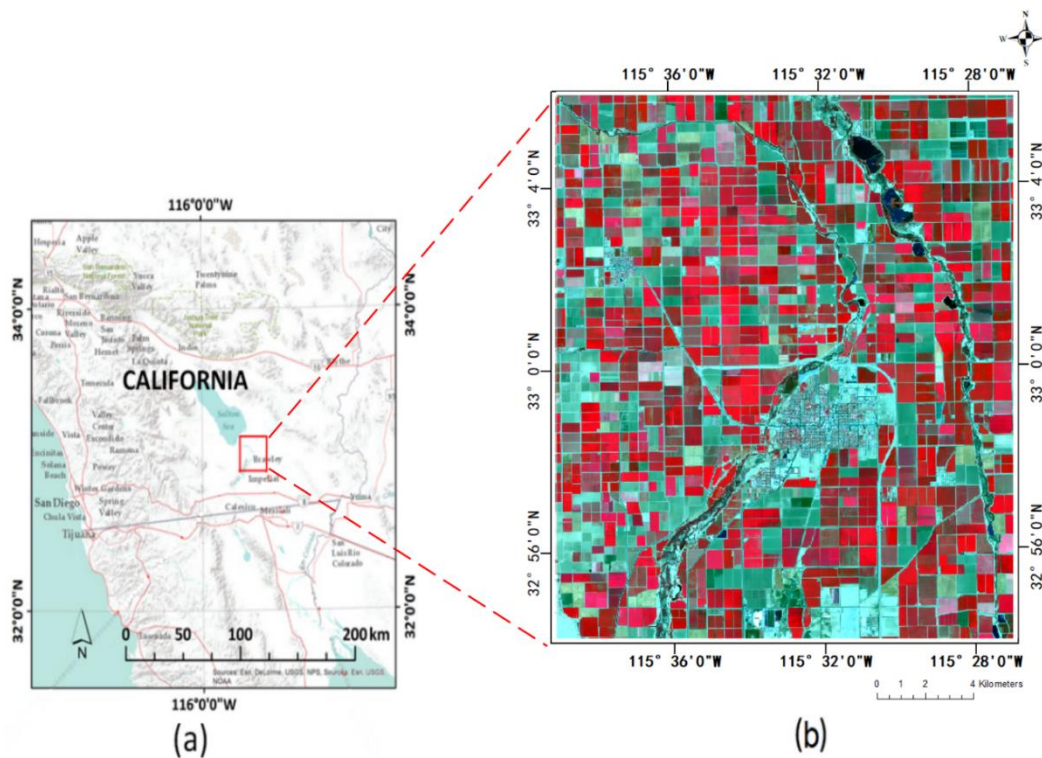


Figure 3 Study area in the Imperial Valley of California, USA (a). The false color combination (band 8 near-infrared, band 4 red, band 3 green) for the study area (May 11, 2017) is depicted in (b).

3.2. Data collection and pre-processing

Sentinel-2 data are used in this study. The data are downloaded from the European Space Agency's (ESA) Sentinel Scientific Data Hub. <https://scihub.copernicus.eu/>. Refer to Table 1, the Sentinel-2 mission has multi-spectral data with 13 bands in the visible, near infrared and short-wave infrared part of the spectrum. The spatial resolution is in 10 m, 20 m and 60 m. Revisiting every 5 days under the same viewing angles. Table 2 shows the dates of the images used in this study. All the data are Sentinel-2 Level 1C, with a cloud cover less than 5%. Only red band (band 4) and the near-infrared band (band 8) of Sentinel-2 data were used. The atmospheric correction was applied using the Sentinel Application Platform (SNAP) v6.0 and the sen2cor plugin v2.5.5 (ESA, 2018b). The Top-Of-Atmosphere (TOA) Level 1C images were processed to surface reflectance Level 2A images which can be directly used to get the Normalized Difference Vegetation Index (NDVI) (Tucker, 1979) time-series data. Resampling was done with the NDVI calculated from Sentinel-2 data from 10m to 30m resolution by using 3×3 window mean filter due to the ground truth data from CropScape was in 30m resolution.

Table 1: Spectral bands for the Sentinel-2 sensors (ESA, 2018)

<i>Sentinel-2 bands</i>	<i>Central wavelength (nm)</i>	<i>Bandwidth (nm)</i>	<i>Spatial resolution (m)</i>
<i>Band 1 - Coastal aerosol</i>	443	20	60
<i>Band 2 - Blue</i>	490	65	10
<i>Band 3 - Green</i>	560	35	10
<i>Band 4 - Red</i>	665	30	10
<i>Band 5 - Vegetation red edge</i>	705	15	20
<i>Band 6 - Vegetation red edge</i>	740	15	20
<i>Band 7 - Vegetation red edge</i>	783	20	20
<i>Band 8 - NIR</i>	842	115	10
<i>Band 8A - Narrow NIR</i>	865	20	20
<i>Band 9 - Water vapour</i>	945	20	60
<i>Band 10 - SWIR-Cirrus</i>	1380	30	60
<i>Band 11 - SWIR</i>	1610	90	20
<i>Band 12 - SWIR</i>	2190	180	20

Table 2: Dates list of the Sentinel-2 time series data used in this research. “No.” represents the number of the time series data.

No.	Date	No.	Date	No.	Date	No.	Date
1	2016-09-03	7	2017-01-31	13	2017-06-20	19	2017-10-08
2	2016-09-23	8	2017-03-02	14	2017-07-10	20	2017-10-23
3	2016-10-03	9	2017-04-21	15	2017-08-14	21	2017-10-28
4	2016-10-13	10	2017-05-01	16	2017-08-29	22	2017-11-22
5	2016-11-22	11	2017-05-11	17	2017-09-18	23	2017-12-07
6	2017-01-01	12	2017-05-21	18	2017-09-23	24	2017-12-22

3.3. Methods

3.3.1. Overview

This research is aimed at finding the most suitable method to the study area’s crops classification among the three DTW-based methods and reducing the computation time of this method. The methods consist of the following steps: (1) data pre-processing to create images of surface reflectance (Level 2A) from TOA (Level 1C)(as mentioned in Section 2.2); (2) data processing, which involves the generation of training and validation samples and the creation of the NDVI time-series; (3) dates choosing, among the 24 Sentinel-2 data listed in Table 2, applying time series classification by using the original DTW method, choose the highest accuracy period as our study period; (4) data classification, where TWDTW, WDDTW, N’DTW are used to classify the target classes; (5) Method selection, choose the most suitable method by assessing the classification accuracies and computation time obtained by the three DTW-based methods. (6) reducing computation time, defining the warping path window and decomposing the patterns are used to reduce the computation time of selected method. (6) Evaluation, for evaluating the classification accuracies and computation time obtained by selected method with the operations of time reducing.

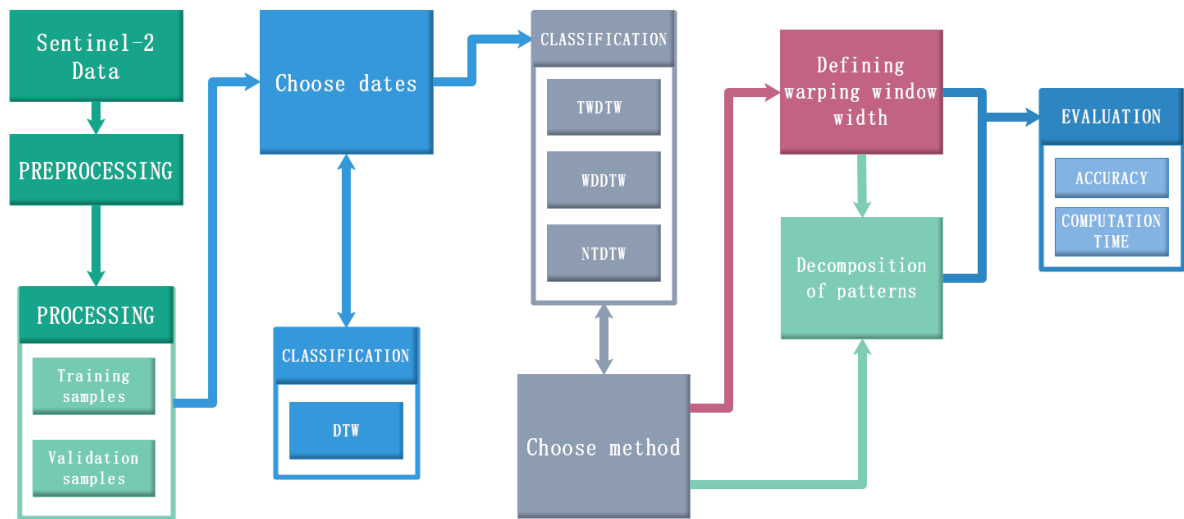


Figure 4 Methods flowchart (DTW represents dynamic time warping; TWDTW represents weighted dynamic time warping; WDDTW represents weighted derivative dynamic time warping; NDTW represents non-isometric transform dynamic time warping)

3.3.2. Training and validation samples

Ground truth data is provided by CropScape - Cropland Data Layer (CDL) for 2017, generated by the United States Department of Agriculture, National Agricultural Statistics Service (Boryan et al., 2011). The spatial resolution of CropScape data is 30 meters. Table 3 shows the first 9 crop types sorted descending by pixel counts. The third column represents the crops' percentage. With the exception of "shrubland" and "developed", the other seven classes were selected for analysis. The numbers of training and validation samples are also listed in Table 3.

The accuracy assessments on CropScape (USDA, 2017b) present the producer and user accuracy (Table 3 Column 6 and 7) of the CDL classification results for California. Only the Alfalfa's Producer Accuracy (PA) is higher than 90%, which means the ground truth data are not reliable. In this study, 50 training samples for each crop type were first plotted their temporal profiles and then compared to the crop calendars (Rabi, 2017; University of California Cooperative Extension, 2017), among 50 samples, the points whose temporal profiles had significantly difference with the others' were removed manually. The validation samples were generated randomly excluding the parcels which had training samples. But, due to the low accuracy of the ground truth data, the validation samples were manually edited. We plotted all the validation samples' temporal profiles and removed the samples who had significant difference with the other samples to make sure the temporal profiles belong to each crop type. In this way, the effects of misclassification on the quality of the validation samples can be alleviated.

Table 3 List of the first 9 crops sorted descending by pixels count; Column 3 is the crops' pixels percentage in this study area; Column 4 and 5 are the numbers of training and validation samples. The numbers in the brackets of training samples are the numbers first selected to choose each crop's pattern. The numbers before the brackets are the final numbers of training samples. Column 6 and 7 are the Producer Accuracy (PA) and User Accuracy (UA) of CropScape classification approach for California derived from the CropScape accuracy assessments (USDA, 2017b).

CLASS_NAME	Pixels count	Percentage (%)	Training	Validation	PA	UA
Alfalfa	155811	33.72%	50	50	91.60%	87.50%
Fallow	64168	13.89%	50	50	80.80%	80.50%
OtherHay	34403	7.45%	50	50	53.60%	65.50%
Shrubland	28455	6.16%	<i>NA</i>	<i>NA</i>	<i>NA</i>	<i>NA</i>
Onions	27854	6.03%	50	50	78.60%	71.40%
Developed	23813	5.15%	<i>NA</i>	<i>NA</i>	<i>NA</i>	<i>NA</i>
Sugarbeets	15997	3.46%	50	50	46.70%	86.00%
WinterWheat	15957	3.45%	50	50	68.10%	72.40%
Lettuce	14531	3.14%	50	50	20.10%	45.10%

3.3.3. Temporal phenological patterns

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:

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad \text{equation (1)}$$

where NIR are data from the Sentinel-2 near-infrared band, and RED are data from the Sentinel-2 visible red band.

NDVI has become one of the most-used indices for studying vegetation phenology (Yan & Roy, 2014). NDVI time series is a good solution to address the issue that describing trends or discrete change events in agriculture land use (Ali, 2009) and it also can reduce the spectral noise caused by illumination conditions, topographic variations or cloud shadows (Huete et al., 2002).

After plotting 18 NDVI time series values of 50 training sample points in each class, we found that the temporal profiles of each class varied, but for the most part shared similarities. However, some points' temporal profiles were very different and were attributed to the mis-classification of the CropScape approach. For example, there were some training points' temporal profiles whose NDVI values were around 0.1 through the whole year. They should be fallow, but the CropScape classified them as other crops. This kind of errors can be easily removed manually. Among the 50 training samples of each class, we manually removed the significant different ones and did the arithmetic averaging for the rest of the training points. The averaging sequences were the patterns of the target 7 classes.

As irrigation was applied in the study area, the area was divided into dense agricultural parcels for the production of cash crops. Different agricultural practices make the temporal phenological patterns of the target crops complex. (El-Gammal et al., 2013) set three levels for different NDVI values. The NDVI values of some samples are below 0 are “No plant cover”; between 0.01 and 0.3 are “Weak plants”; above 0.31 are “Healthy plants”. Fallow is a class with values of NDVI around 0.1 through the year. Alfalfa and other hay (non-alfalfa) may be planted and harvested periodically throughout the whole year. So, their patterns show the maximum irregularity among the other crop patterns. Sugar beets has a trapezoid like temporal profile. The NDVI values start increasing in the middle of October 2016. Before decreasing in May 2017, the sugar beets stayed “Healthy plants” for 7 months. Winter wheat and onions have one highest NDVI value in their temporal profiles. For winter wheat, the values increased in November and got the peak in February. The increasing of onions values had two months delayed and got the peak in April. After June, both of them became bare soil. Lettuce has two peaks, in Dec-Jan and May, followed by bare soil afterwards. Figure 8(b) illustrates the phenological patterns of the target crops using Sentinel-2 NDVI time series for 2017.

3.3.4. Classification

After the 10 m to 30 m resolution resampling, the size of NDVI time series images was reduced from (2190, 1899) to (730, 633). There were 462090 pixels to be classified into 7 target classes. However, the DTW and DTW-based methods are all time-consuming (Belgiu & Csillik, 2018; Górecki & Łuczak, 2014; Jeong et al., 2011; Maus et al., 2016; Petitjean et al., 2012). A subset (Figure 5) of the study area was defined to accelerate the testing of DTW and DTW-based methods’ performance. The size of the subset is (300, 300) with 90,000 pixels. The crop types in the subset had the similar proportion with the initial study area. The classification was first carried out in the selected subset. The distance calculations between each pixel’s sequence (i.e. NDVI time series data) and 7 temporal patterns was performed in this study. Among the 7 results, the pixel is assigned to the class which had the smallest distance value. In this study, 3 DTW-based methods (TWDTW, WDDTW, NTDTW) of distance calculation were applied on the classification of Sentinel-2 NDVI time-series data and the original DTW method was also applied as comparison. After the three DTW-based methods applied on the subset of study area, we chose the highest accuracy one which was the TWDTW to apply on the whole study area (2190,1899). The warping path window was defined on the TWDTW distance calculation to reduce the computation time. We also tried to decompose the temporal profiles of the patterns to decrease the computation time.

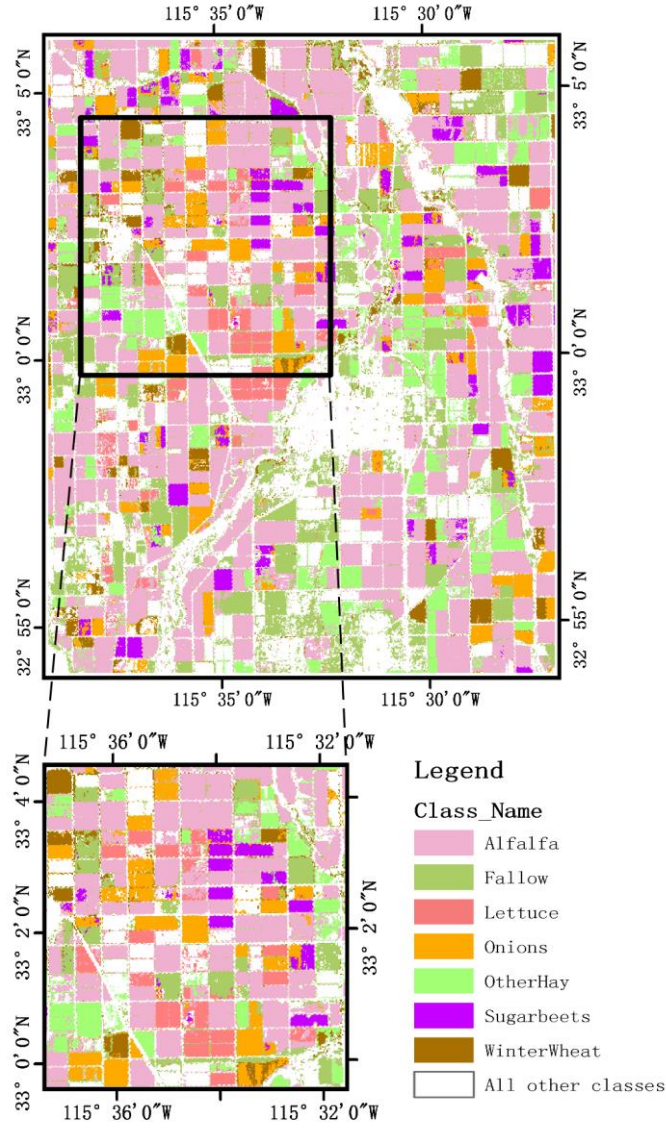


Figure 5 Study area (upper) and (300, 300) subset (below) CropScape land use map. The legend adopted from (USDA, 2018)

3.3.5. DTW

DTW is an algorithm which is based on Euclidean distance for similarity measurement between two sequences. The two time series sequences $\mathbf{U} = \{u_1, u_2, \dots, u_n\}$ and $\mathbf{V} = \{v_1, v_2, \dots, v_m\}$ have lengths n and m . A $n \times m$ matrix $\mathbf{D}_{base} = (d_{base}(u_i, v_j))_{n \times m}$ is used to store the Euclidean distance between $u_i \in \mathbf{U} \forall i = 1, 2, \dots, n$ and $v_j \in \mathbf{V} \forall j = 1, 2, \dots, m$.

$$d_{base(i,j)} = |u_i - v_j| \quad \text{equation (2)}$$

The DTW distance \mathbf{D} is computed by a recursive sum of the minimal distances, such that,

$$d_{i,j} = d_{base(i,j)} + \min\{d_{i-1,j}, d_{i-1,j-1}, d_{i,j-1}\} \quad \text{equation (3)}$$

There are boundary conditions of the \mathbf{D} :

$$d_{i,j} = \begin{cases} d_{base(i,j)}, i = 1, j = 1 \\ \sum_{k=1}^i d_{base(k,j)}, 1 < i \leq n, j = 1 \\ \sum_{k=1}^j d_{base(i,k)}, i = 1, 1 < j \leq m \end{cases} \quad \text{equation (4)}$$

Figure 6 illustrated the DTW distance matrix. Each cell in the matrix saved the DTW distance value between two sequences. The red cells which represented the lowest DTW distance values composed the warping path. The farther away is the warping path from the diagonal, the more points on one sequence will be matched to one point on the other sequence.

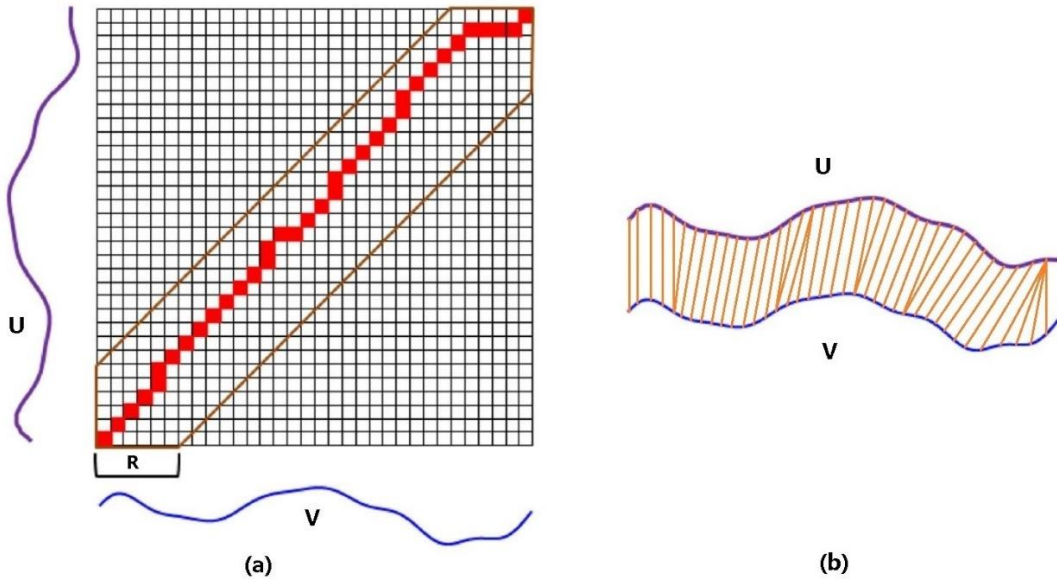


Figure 6 DTW distance matrix and Warping path window (Jeong et al., 2011). U and V are the two sequences to be calculated for DTW distance. (a) shows the DTW distance matrix of U and V. The area circled by brown line is the warping path window. R represents the warping path window width. (b) shows the alignment of U and V based on DTW.

3.3.6. TWDTW

In the Time-Weighted Dynamic Time Warping (TWDTW) method (Maus et al., 2016), a weight ω is given to the d_{base} ,

$$d_{base(i,j)} = \omega_{i,j} |u_i - v_j| \quad \text{equation (5)}$$

The WDTW distance **WD** is such that,

$$wd_{i,j} = \omega_{i,j} |u_i - v_j| + \min\{wd_{i-1,j}, wd_{i-1,j-1}, wd_{i,j-1}\} \quad \text{equation (6)}$$

The weight ω is defined as below called the modified logistic weighted function (MLWF),

$$\omega_{i,j} = \frac{1}{1 + e^{-\alpha(g(t_i, t_j) - \beta)}} \quad \text{equation (7)}$$

The $g(t_i, t_j)$ is the elapsed time in days between the dates t_i in the pattern \mathbf{U} and t_j in \mathbf{V} . steepness a controls the level of penalization for the points with larger phase difference. β is the midpoint.

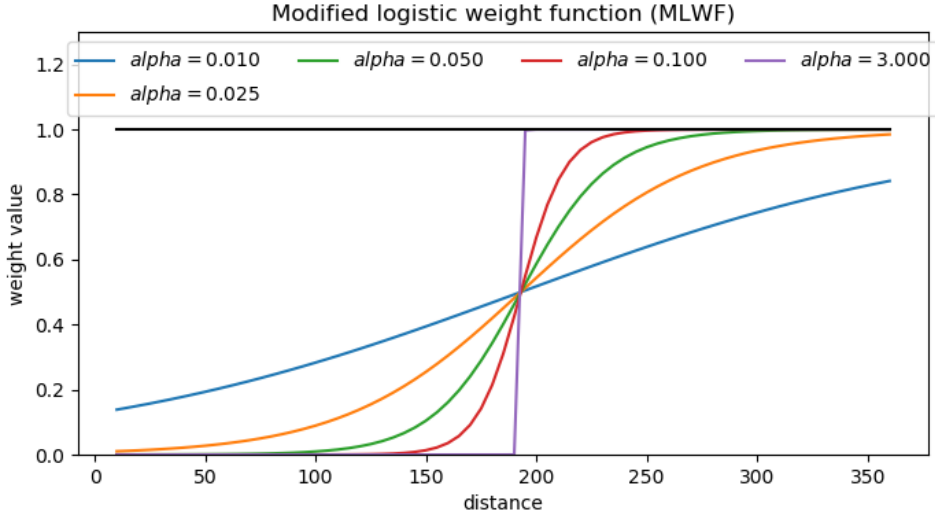


Figure 7 MLWF with different α .

The value a could range from zero to infinity, there are four cases with different a as Figure 7 shown: (1) Constant weight: when $a=0$, all points are given the same weight; (2) Linear weight: when $a=0.01$, the weight value is nearly a linear relationship; (3) Sigmoid weight: different value of a can achieve different sigmoid patterns. Figure 7 shows the sigmoid pattern when $a=0.025, 0.05, 0.1$. (4) Two distinct weights: when $a=3$, the first one-half is given one weight value, the second one-half is given another weight value.

In the time series classification, the sigmoid weight is used to give different penalty to different time warps. There will be a low penalty for small time warps and significant cost for large time warps. But the a and β must be tuned to suit different datasets. In this study, after tuning, we set the $a=0.025, \beta=193$.

3.3.7. WDDTW

Weighted derivative dynamic time warping (WDDTW) (Jeong et al., 2011) is based on the TWDTW and imposes the idea of derivative dynamic time warping (DDTW). The DDTW transforms the original points into higher level features, which contain the shape information of a sequence. The estimate equation for transforming data point u_i in the sequence \mathbf{U} is given by,

$$d_i^u = \frac{(u_i - u_{i-1}) + ((u_{i+1} - u_{i-1}) / 2)}{2}, 1 < i < n \quad \text{equation (8)}$$

where n is the length of sequence \mathbf{U} . $d_1^u = d_1^u$ and $d_n^u = d_{n-1}^u$.

Then the weighted version of DDTW (WDDTW) is given as follows:

$$wdd_{i,j} = \omega_{i,j} |d_i^u - d_j^v| + \min\{wdd_{i-1,j}, wdd_{i-1,j-1}, wdd_{i,j-1}\} \quad \text{equation (9)}$$

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

The estimate d_i^u is simply the average of the slope of the line through the test point and its left neighbor, and the slope of the line through the left neighbor and the right neighbor. This method uses the shape features instead of original value to alleviate the confusion caused by local differences (Figure 1) and weights.

3.3.8. NDTW

In case of non-isometric transform dynamic time warping (NTDTW), there are three transforms used for the DTW distance measure. For the time series $U = \{u_i : i = 1, 2, \dots, n\}$ has a transform

$$T = \{u_k : k = 1, 2, \dots, n\}.$$

$$\text{Cosine transform: } t_{k_u} = \sum_{i=1}^n u_i \cos\left[\frac{\pi}{n} \left(i - \frac{1}{2}\right) (k_u - 1)\right] \quad \text{equation (10)}$$

$$\text{Sine transform: } t_{k_u} = \sum_{i=1}^n u_i \sin\left[\frac{\pi}{n} \left(i - \frac{1}{2}\right) k_u\right] \quad \text{equation (11)}$$

$$\text{Hilbert transform: } t_{k_u} = \sum_{\substack{i=1 \\ i \neq k}}^n \frac{u_i}{k_u - i} \quad \text{equation (12)}$$

The NTDTW distance is defined as:

$$td_{i,j} = (1 - \theta)d_{i,j} + \theta d_{k_u, k_v}$$

where $d_{i,j}$ is the DTW distance between U and V , d_{k_u, k_v} is the DTW distance between transform U and transform V .

The mathematical transforms are popularly used in the classification of time series to extract higher order features. This method adds transforms of the sequences to the DTW distance and uses a parameter θ to control the effect of the transforms. Then the final distance will include the features of higher order. The parameter θ has to be tuned.

3.3.9. Defining the warping path window width

Sakoe (1978) proposed the Sakoe-Chiba Band to constrain the calculation of DTW distance to a specific warping path window. The Sakoe-Chiba Band not only reduces the computation time of the DTW distance calculation, but also improve the accuracy (Sakoe, 1978).

There is a warping path $w_k = (i, j)_k$ of two sequences whose length are i and j , the constraint of the warping path is such that $j - R_i \leq i \leq j + R_i$, where R_i is a term defining the allowed range of warping, for a given point in a sequence. In the case of the Sakoe-Chiba Band, R is independent of i .

$$R_i = d, 0 \leq d \leq m, 1 \leq i \leq m$$

where R_i is the height above the diagonal in the y-direction, as well as the width to the right of the diagonal in the x-direction. Note that $|R|=m$, and the above definition forces R to be symmetric.

As an example, a Sakeo-Chiba Band of overall width of 11 (width 5 strictly above and to the right of the diagonal) with the definition

$$R_i = \begin{cases} 5 & 1 \leq i \leq m-5 \\ m-i & m-5 < i \leq m \end{cases} \quad \text{equation (13)}$$

In Figure 6(a), the area circled by the brown line is the warping path window. R is the width of the window. Warping path window constraints the DTW distance calculation and make sure that one point on one sequence will not be matched to more than R numbers of points. In terms of temporal aspect, the warping path window limits the warping which will not happen between points which had more than R numbers of intervals. The effect of the warping path window and the weight in TWDTW method are alike. They all try to constrain the DTW distance calculation to prevent the points on one sequence match points which are too far away on the other sequence. For TWDTW, the set of warping path window width should not interfere with the effect of the weight. The intervals of selected time series' dates in this study is not equivalent, with a range from 10-40 days. So, the maximum warping days controlled by the warping path window are $10R-40R$. In MLWF of TWDTW, the maximum warping days is decided by the value of β . If we do not want the defining of warping path window to affect the results of TWDTW method, we should make sure $10R-40R$ is larger than β .

3.3.10. Decomposition of phenological patterns

For the patterns of target classes, there were many periods when the temporal profiles of the crops overlapped. Most of the overlaps happened when the soil was bare. Except Alfalfa and other hay, the other 5 classes were overlapping from Sep to Oct 2016 and from Aug to Sep 2017. The DTW distance between any two of these 5 classes during these periods would be very small. In the classification, we calculated the DTW distance between each pixel and 7 classes' patterns. In the bare soil periods of the 5 classes, the short DTW distance had no contribution to the distinguishing of these 5 classes. The discrimination between target crops relies on the other periods. Therefore, we reduce some classes' calculation periods. For fallow, due to the fact that it has low value throughout the whole year, we only kept March-June part where the fallow had distinct values from the other 6 classes. For onions and winter wheat, the period from December to June was kept. Their different peak times can be used to distinguish them. Lettuce had two peaks in the temporal profiles. We kept the two peaks and removed the bare soil periods. Sugar beets had a trapezoid like temporal profile, but we notice that if we remove the bare soil periods of sugar beets, the rest of the profile would be similar to alfalfa. Therefore, the bare soil periods are useful for distinguishing sugar beets and alfalfa. Consequently, we kept the whole temporal profile of sugar beets. Alfalfa and other hay fluctuate throughout the entire year. We kept the whole temporal profiles of them.

4. RESULTS AND DISCUSSION

4.1. Results

All the DTW and DTW-based methods analysis in this study were run on a configuration with 4 cores 3.30 GHz and 8-GB memory in PyCharm 2018.2.4 64-bit on Windows 10 64-bit professional version. The code of this thesis can be found on <https://zenodo.org/record/2594722#.XIuC4ShKhPY>. The DOI: 10.5281/zenodo.2594722. In the programming, we used “numpy” library for processing the matrices. “math” for calculating the Euclidean distance. The main program “test.py” has the ability to call all the function we need in this study. DTW and DTW-based methods are all coded as functions. The DTW and TWDTW with warping path window are coded as independent functions. The decomposition patterns of DTW and TWDTW are coded as independent functions.

4.1.1. Phenological patterns decision

At the beginning of this research, we downloaded only the Sentinel-2 data of 2017 from January to December (total 18 dates). From the training samples, we plotted the temporal profiles of the target seven crops (Figure 8(a)). We changed the dates and plotted the temporal profiles of targets crops from September 2016 to September 2017 as Figure 8(b) shown. Figure 8(c) shows the dates of the reduced temporal profiles from September 2016 to June 2017.

The original DTW classifications were applied on these three periods, the accuracies of these three periods' classifications were shown in Table 4. Among the three periods, Sep2016-Sep2017 period achieved the highest overall accuracy, which was 57.4%. The overall accuracies yielded for Jan2017-Dec2017 period and Sep2016-June2017 period were 52% and 54% respectively. The computation time of Jan2017-Dec2017 and Sep2016-Sep2017 were similar, due to the fact that there were 18 dates in their time series data. For Sep2016-June2017 period, there were 13 dates in its time series, which reduced the computation time.

The classification results are presented in Table 5. The best classified class was the fallow class with a PA and UA of 100%. Alfalfa and OtherHay class were confused with each other. Many OtherHay class validation samples were classified as Alfalfa. This confusion decreased the UA of Alfalfa class and the PA of OtherHay class. The high overlap between Onions, Lettuce and WinterWheat class lead to the low PAs and UAs of these three classes. Some Sugarbeets validation samples were misclassified as Onions and WinterWheat, which decreased the PA of Sugarbeets to 64%.

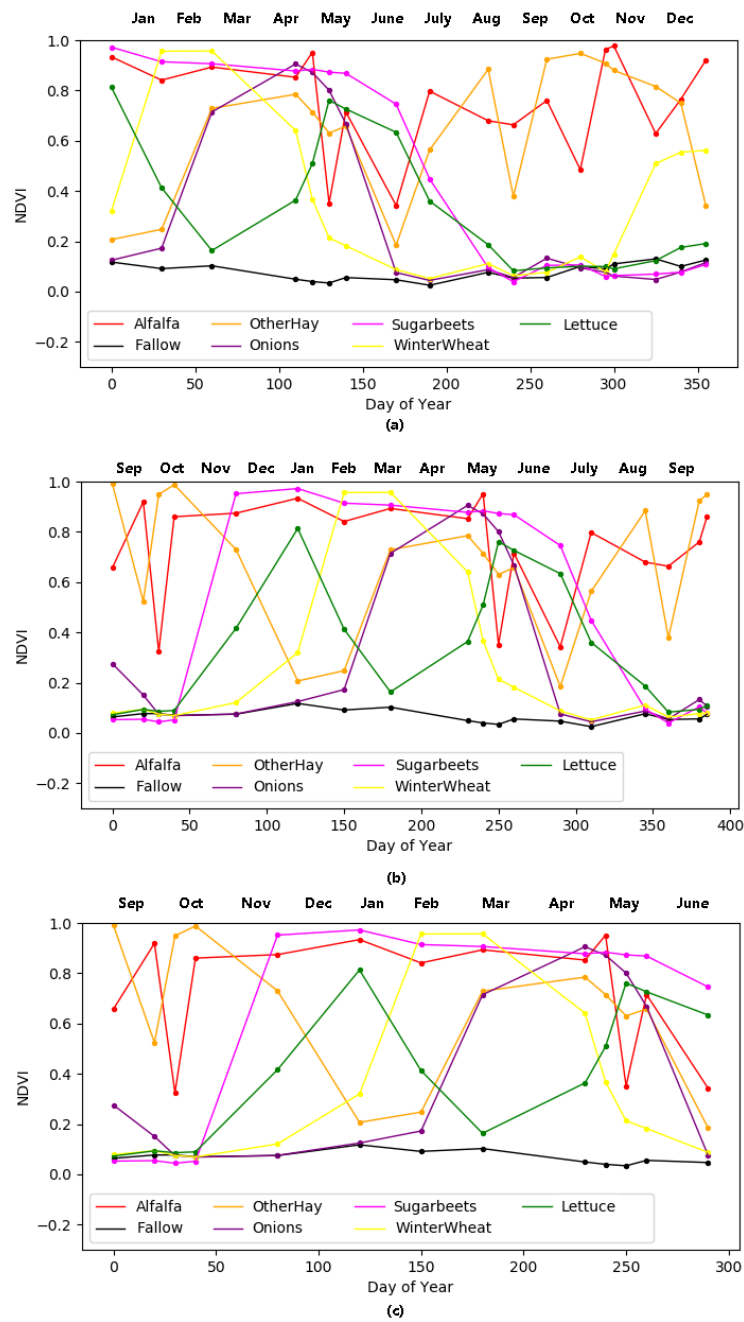


Figure 8 Temporal profiles of different dates selection of Sentinel-2 time series data. (a) shows the 7 crops' temporal profiles from January 2017 to December 2017 (total 355 days). (b) shows the 7 crops' temporal profiles from September 2016 to September 2017 (total 385 days). (c) shows the 7 crops' temporal profiles from September 2016 to June 2017 (total 290 days).

Table 4 Overall accuracies and computation time of three periods' (January 2017-December 2017, September 2016-September 2017, September 2016-June 2017) DTW classifications.

Periods	Overall Accuracy	Computation time
Jan2017-Dec2017	52%	7261s
Sep2016-Sep2017	57.4%	7253s
Sep2016-June2017	54%	4689s

Table 5 Classification accuracy of DTW method from September 2016 to September 2017.

<i>DTW</i>									
<i>Croptypes</i>	Alfalfa	Fallow	Onions	OtherHay	Lettuce	WinterWheat	Sugarbeets	Total	PA
<i>Alfalfa</i>	42	0	1	0	3	0	4	50	0.84
<i>Fallow</i>	0	50	0	0	0	0	0	50	1
<i>Onions</i>	0	0	22	0	12	16	0	50	0.44
<i>OtherHay</i>	34	0	0	16	0	0	0	50	0.32
<i>Lettuce</i>	0	0	9	0	21	19	1	50	0.42
<i>WinterWheat</i>	0	0	32	0	0	18	0	50	0.36
<i>Sugarbeets</i>	0	0	11	0	0	7	32	50	0.64
<i>Total</i>	76	50	75	16	36	60	37	350	
<i>UA</i>	0.553	1	0.293	1	0.583	0.3	0.865		0.574

4.1.2. Classification experiments on three DTW-based methods

Given the classification results, we chose September 2016 to September 2017 as our study period. Three DTW-based method classifications were performed on this period. Table 6 depicts the results of the 3 DTW-based method classifications. The two tuneable parameters *alpha* and *beta* of the TWDTW method were set to 0.025 and 193 through trail-and-error. The computation time of TWDTW classification for 7 target crops was 16398 seconds. For the WDDTW method, the *alpha* parameter was tuned to 0.4. The computation time of the WDDTW classification was 9385 seconds. In case of the NTDTW method, a value of 0.94 was used for the *alpha* parameter. The computation time of NTDTW classification was 9313 seconds. In case of the TWDTW method, the UAs and PAs for Alfalfa, Fallow, OtherHay, and Sugarbeets achieved high values (higher than 90%). The PA for WinterWheat class was 98%, while some Onions fields were misclassified as WinterWheat which contributed to the reduction of the UA of WinterWheat to 81.7% and the PA of Onions to 76%. The UA of Onions class was 67.9%, due to confusion with the Lettuce class.

In case of the WDDTW, Fallow class achieved a PA and UA of 100%. Sugarbeets class yielded high PA and UA, namely 80% and 87% respectively. Alfalfa class had a PA of 90%, while OtherHay class was confused with Alfalfa which contributed to the reduction of the UA of Alfalfa to 61.6% and the PA of OtherHay to 42%. As no other class was classified as OtherHay class, OtherHay achieved a UA of 100%. Onions, Lettuce and WinterWheat presented a high confusion with each other which led to relatively low PAs and UAs values.

For NTDTW, Fallow and Sugarbeets got high PAs and UAs. The confusion of Onions, Lettuce and WinterWheat caused low PAs and UAs values for these classes. The PAs of Onions and WinterWheat were only 20% and 30% respectively. Many Onions and Lettuces were classified as WinterWheat which decreased the UA of WinterWheat to 23.8%. OtherHay class was confused with Alfalfa. Therefore, the PA of OtherHay was 52%.

Table 6 Classification accuracy assessments of TWDTW, WDDTW and NDTW methods. (a) TWDTW, (b) WDDTW (c) NDTW. PA represents producer's accuracy, UA represents user's accuracy.

TWDTW

<i>Croptypes</i>	Alfalfa	Fallow	Onions	OtherHay	Lettuce	WinterWheat	Sugarbeets	Total	PA
<i>Alfalfa</i>	47	0	1	0	0	0	2	50	0.94
<i>Fallow</i>	0	50	0	0	0	0	0	50	1
<i>Onions</i>	0	2	38	0	0	10	0	50	0.76
<i>OtherHay</i>	5	0	0	45	0	0	0	50	0.9
<i>Lettuce</i>	0	0	16	0	30	1	3	50	0.6
<i>WinterWheat</i>	0	0	1	0	0	49	0	50	0.98
<i>Sugarbeets</i>	0	0	0	0	0	0	50	50	1
<i>Total</i>	52	52	56	45	30	60	55	350	
<i>UA</i>	0.904	0.962	0.679	1	1	0.817	0.909		0.883

(a)

WDDTW

<i>Croptypes</i>	Alfalfa	Fallow	Onions	OtherHay	Lettuce	WinterWheat	Sugarbeets	Total	PA
<i>Alfalfa</i>	45	0	1	0	1	0	3	50	0.9
<i>Fallow</i>	0	50	0	0	0	0	0	50	1
<i>Onions</i>	0	0	18	0	12	20	0	50	0.36
<i>OtherHay</i>	28	0	0	21	1	0	0	50	0.42
<i>Lettuce</i>	0	0	5	0	32	10	3	50	0.64
<i>WinterWheat</i>	0	0	19	0	0	31	0	50	0.62
<i>Sugarbeets</i>	0	0	2	0	0	8	40	50	0.8
<i>Total</i>	73	50	45	21	46	69	46	350	
<i>UA</i>	0.616	1	0.4	1	0.696	0.449	0.87		0.667

(b)

NTDTW

<i>Croptypes</i>	Alfalfa	Fallow	Onions	OtherHay	Lettuce	WinterWheat	Sugarbeets	Total	PA
<i>Alfalfa</i>	41	0	0	7	0	0	2	50	0.82
<i>Fallow</i>	0	47	2	0	0	1	0	50	0.94
<i>Onions</i>	0	1	10	0	12	27	0	50	0.2
<i>OtherHay</i>	22	0	2	26	0	0	0	50	0.52
<i>Lettuce</i>	0	0	4	0	26	20	0	50	0.52
<i>WinterWheat</i>	2	0	6	27	0	15	0	50	0.3
<i>Sugarbeets</i>	0	0	0	0	0	0	50	50	1
<i>Total</i>	65	48	24	60	38	63	52	350	
<i>UA</i>	0.63	0.979	0.417	0.433	0.684	0.238	0.962		0.614

(c)

4.1.3. Defining the warping path window

The defining of warping path window is to reduce the computation time of the DTW-based methods. Among the 3 DTW-based methods, the TWDTW method was chosen to define the warping path window as the result of the highest overall accuracy and longest computation time. The definition of the warping path window of original DTW method was also considered as comparison. Table 7 showed the classification results of the DTW and TWDTW methods obtained by defining different widths of the warping path window.

In case of the DTW, the overall accuracy increased from 57.4% to 86.3% and the computation time decreased from 5438 seconds to 1467 seconds, with the warping path window width decreasing. For TWDTW, we made subtraction between the classification result of the original TWDTW method and the classification results of TWDTW with different warping path window widths. The pixel differences column shows the pixel numbers of each width classification result difference with the original TWDTW method. The overall accuracy was 88.3%, and did not change much when the warping window width was changed from 5 to 1. From the width 5 to 4, the classification results did not change at all. For width 3, 2 and 1, the pixel numbers were 53, 370 and 3507 respectively. The computation time decreased with the lower warping path window width from 8607 seconds to 6100 seconds. Figure 9 shows the PAs and UAs of DTW and TWDTW classification.

Table 7 The Overall Accuracy (OA) and computation time (Time) of DTW and TWDTW methods with different warping path window widths. The “pixel differences” column means the pixel numbers of each warping path window width classification result difference with the classification result of the original TWDTW.

Warping path window width	DTW		TWDTW		
	OA	Time	OA	Time	Pixel differences
5	0.574	5438s	0.883	8607s	0
4	0.617	4687s	0.883	7986s	0
3	0.68	2389s	0.883	7132s	53
2	0.734	2212s	0.883	6842s	370
1	0.863	1467s	0.883	6100s	3507

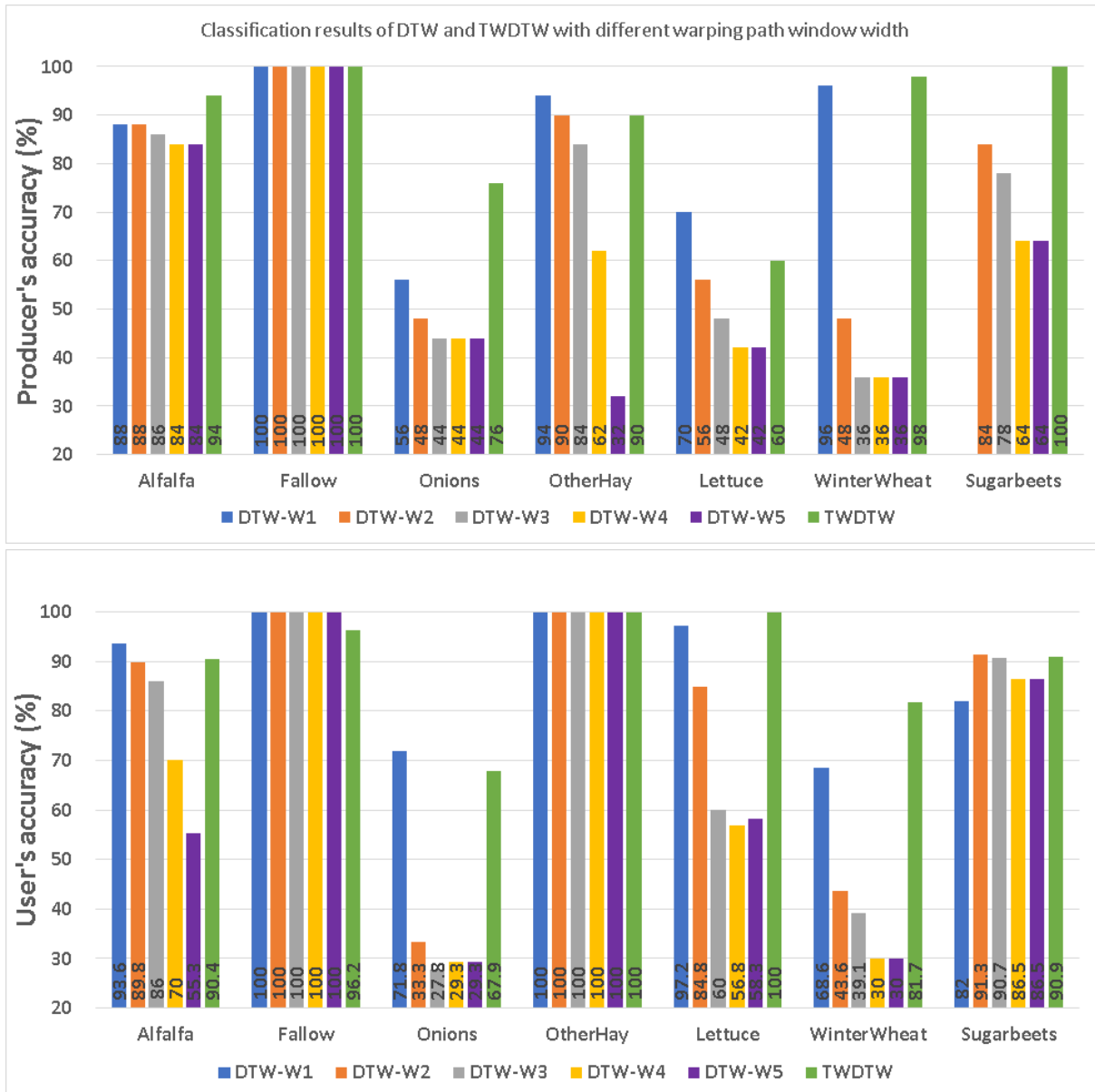


Figure 9 Classification results of DTW with warping path window width from 5 to 1 and TWDTW with warping path window.

4.1.4. Decomposition of phenological patterns

A knowledge-based decomposition of phenological was done on the period September 2016 to September 2017. Classifications were implemented on pattern decomposition DTW (PD-DTW), pattern decomposition DTW with warping path window (PD-DTW-W), pattern decomposition TWDTW (PD-TWDTW), pattern decomposition TWDTW with warping path window (PD-TWDTW-W). We got the classification results of the 4 methods mentioned above are displayed in Figure 10. The overall accuracies and computation time of these 4 methods are listed in Table 8.

In case of the PD-DTW and PD-DTW-W, we obtained an overall accuracy of 68.3% and 74.3%, computation time 2385s and 1992s respectively. The Fallow class obtained PAs of 98%. The OtherHay achieved UAs of 100%. The WinterWheat class achieved relatively low PAs of 30%. The Sugarbeets class

obtained a UA of 40% for PD-DTW and 45.3% for PD-DTW-W, which were relatively low. For PD-DTW, some OtherHay agricultural fields were classified as Alfalfa class which reduced the PA of OtherHay to 60%. The confusion of Sugarbeets, Onions, Lettuce and WinterWheat led to the relatively low PAs and UAs values of these classes.

In case of the PD-TWDTW and PD-TWDTW-W, the overall accuracies were 86.5% and 89.1%, the computation time were 5291s and 4029s. The Fallow class obtained PAs of 100%. The OtherHay achieved UAs of 100%. For PD-TWDTW, except for the Lettuce class which obtained a PA of 70% and the Onions class which obtained a UA of 68.3%, the other classes obtained PAs and UAs of higher 70%. For PD-TWDTW-W, except for the Lettuce class that achieved a PA of 70% and Fallow class which had a UA of 72.5%, the other classes achieved PAs and UAs of higher 80%.

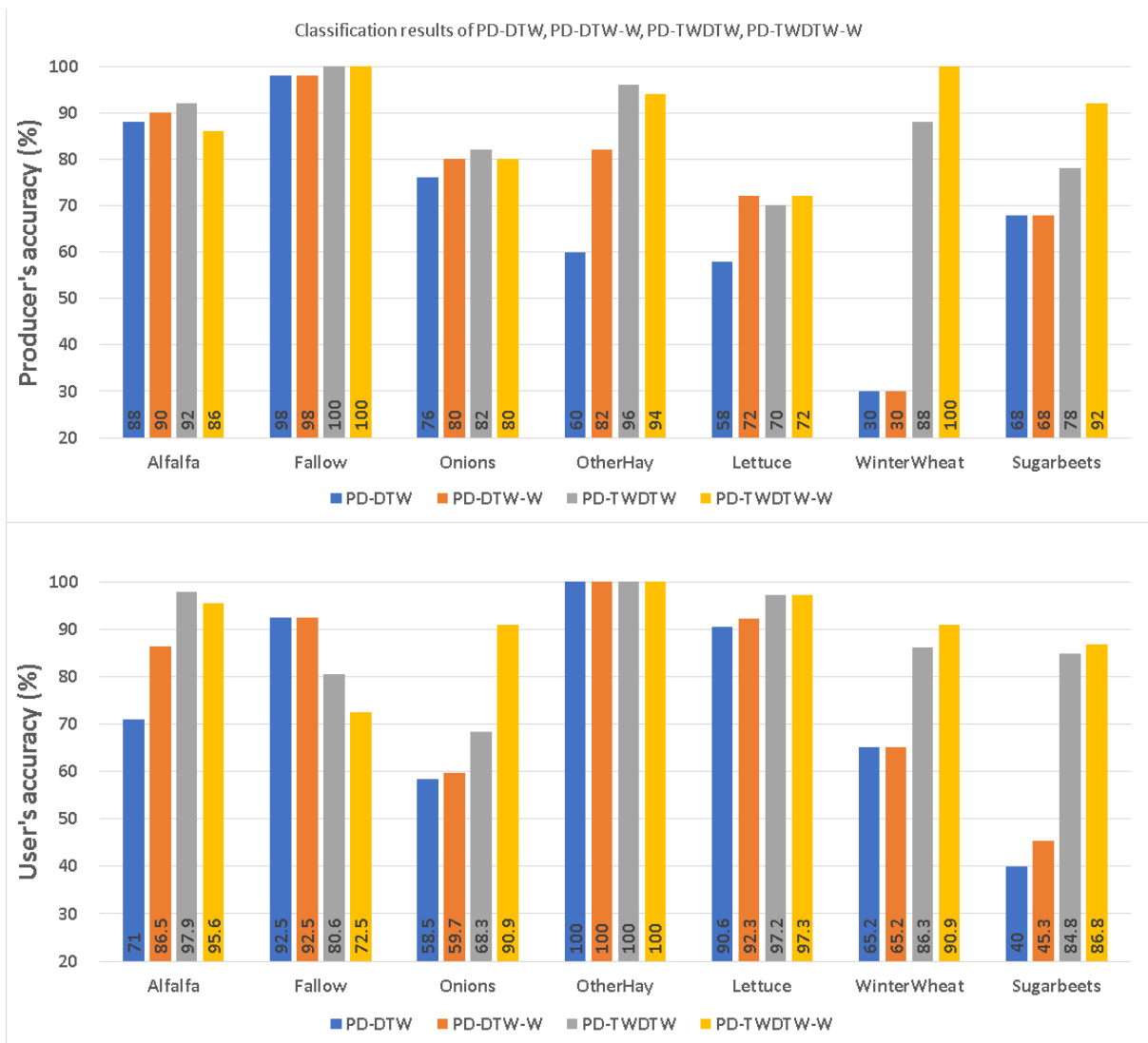
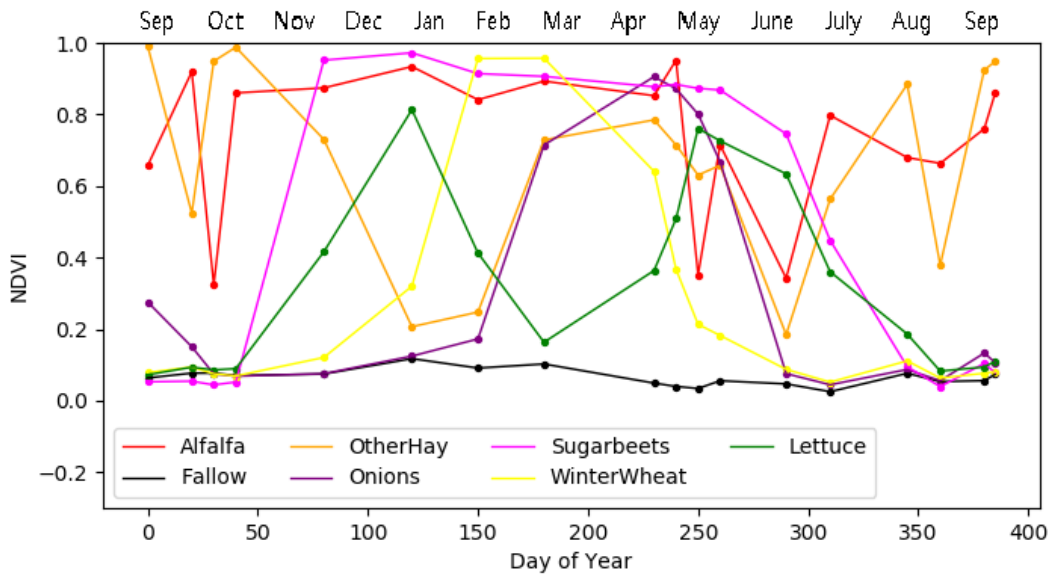


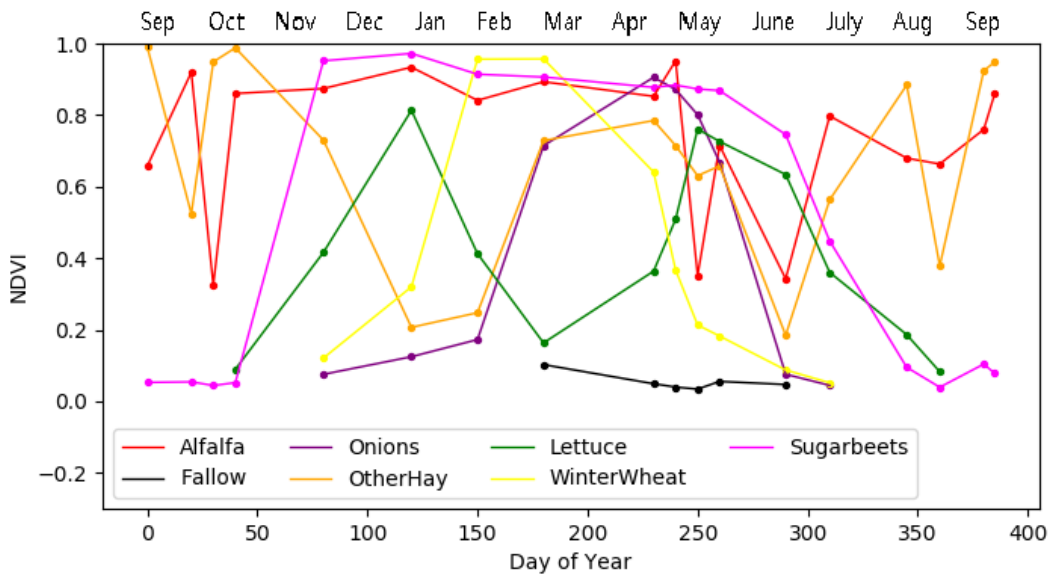
Figure 10 Classification results of pattern decomposition DTW (PD-DTW), pattern decomposition DTW with warping path window (PD-DTW-W), pattern decomposition TWDTW (PD-TWDTW), pattern decomposition TWDTW with warping path window (PD-TWDTW-W).

Table 8 Overall accuracy results and computation time of DTW and TWDTW with and without patterns decomposition classification. PD-DTW, PD-DTW-W, PD-TWDTW, PD-TWDTW-W represent pattern decomposition DTW, pattern decomposition DTW with warping path window, pattern decomposition TWDTW, pattern decomposition TWDTW with warping path window.

Methods	Overall accuracy	Computation time
PD-DTW	0.683	2385s
PD-DTW-W	0.743	1992s
PD-TWDTW	0.865	5291s
PD-TWDTW-W	0.891	4029s



(a)



(b)

Figure 11 Comparison of original patterns and decomposed patterns

4.2. Discussion

4.2.1. Selection of the time series interval

Sentinel-2 provides 10 m optical images every 5 days. Dense time series data from Sentinel-2 contribute significantly to remote sensing applications in agriculture monitoring. Yet, these data come along some challenges that need to be carefully addressed. One of the challenge is how to handle the irregularly-spaced time series data (Roerink et al., 2000). The irregularly-spaced time series data are caused by removing cloud contaminated observations from the 5-day time series data. All of the DTW and DTW-based methods were to measure the similarity between phenological patterns and the time series data. The selection of the time series interval from the irregularly-spaced time series data is of vital importance to the quality of extracting phenological patterns.

At the beginning of this research, we planned to use the crops' temporal profiles of the year 2017 for the classification purpose. Given the fact that the study area was in the desert region, we were not familiar with the phenological cycles of crops of interest. Through visual inspection of the temporal profiles from January 2017 to December 2017, we found out that the NDVI values of the Sugarbeets and Lettuce classes had high value at the beginning. Therefore, we concluded that these two crops start to turn green before January 2017. To get the entire phenological cycles of Sugarbeets and Lettuce, we referred to the crops calendar (Rabi, 2017; University of California Cooperative Extension, 2017) and shifted the study period to September 2016 – September 2017. The temporal profiles of the target 7 crops were shown in Figure 8(b). It is reasonable that the accuracy obtained for the Jan-to-Dec period (namely 52%) is lower than those obtained for the Sept-to-Sept period (namely 57.4%). This happens because most of the crops reached the greenness phase during the October 2016 and their senescence phase around September 2017. The complete growing season was covered by the Sept-to-Sept time series. Furthermore, the Jan-to-Dec pattern not only cut the phenological cycles, but might also contain the crops plantation information of the next year. As different crops can be grown in the same field next year, the classification results might be affected by the selection of the time series interval. Consequently, the time series interval selection was September 2016 to September 2017.

Bégué et al. (2018) claimed that expert knowledge on the local agricultural systems, in particular the knowledge of crop types and crop calendars, plays an important role in the agriculture applications of satellite data. The time series data contains spectral and temporal variability which is caused by the environment, the cropping system and the quality of images (Bégué et al., 2018). Without the local knowledge, it is hard to achieve a high accuracy in extracting phenological patterns and the DTW-based classifications.

4.2.2. Assessment of three DTW-based methods

As the DTW method's classification accuracy shown, 57.4% is not a satisfying result for crops classification accuracy. The three DTW-based methods (TWDTW, WDDTW, NDTW) have their advantages in

classification. One of the objectives of this study was to assess the performance of these three DTW-based methods. We tested these three methods on our study area and got the classification results as Table 6 shown.

4.2.2.1. TWDTW

Among the three DTW-based methods, TWDTW got the highest overall accuracy. The main reason is the TWDTW gives the similarity measurement between the pattern and the test sequence a time constraint. This time constraint limits the point on pattern matches the point which has a large time lag on test sequence. For example, there were many testing samples of WinterWheat class which were classified as Onions class in the DTW method's confusion matrix. However, in TWDTW method, only 1 of 50 WinterWheat testing samples were classified as Onions class. From the temporal profiles shown in Figure 8(b), the patterns of WinterWheat and Onions have a similar profile, which means the durations and amplitudes of their NDVI curves were similar. The difference is that the WinterWheat's NDVI values started to increase, reached its peak and decreased to low value one month earlier than the Onions class. The DTW method failed to consider the time lag between these two classes. The TWDTW method used Modified Logistic Weight Function (MLWF) to define a weight for the distance between two points which have a time lag. The distance between these two points becomes larger with an increasing time lag. Due to the defined weight, TWDTW improved the overall accuracy of classification from 57.4% to 88.3%. In the study of Belgiu & Csillik's (2018) study, pixel-based TWDTW was applied on three study areas. The overall accuracies of TWDTW were 92.14%, 86.78% and 66.34% respectively. It should be mentioned that for the study area 3 in Belgiu & Csillik's (2018) study, we found that the low PA and UA for crop classes was not only from the high intra-class spectral heterogeneity, errors in the patterns, but also from the validation samples. Referring to the accuracy assessments on CropScape (USDA, 2017b), there could be errors in the validation samples. Consequently, in this study, we manually edit the validation samples by removing the samples who had significant difference with the others. In Maus et al.'s (2016) study, TWDTW method was used to classify Single cropping, Double cropping, Forest and Pasture. The Double cropping obtained a PA and UA of 90.43% and 92.04%. The single cropping achieved a PA and UA of 84.85% and 75% respectively. The Lettuce class as double cropping in this study achieved higher PA of 100% and lower UA of 60%, the WinterWheat as single cropping obtained higher PA and UA of 98% and 81.7%.

Despite its clear advantages, the TWDTW presents also some shortcomings that worth being mentioned. The flexible agricultural practices in this study area made the same crop type's temporal profiles to have different planting and harvesting time (Bégué et al., 2018). These time differences not only affect the quality of the training samples but also hinder the classification results. In case of the Onions class, there were some samples whose temporal profiles were delayed than the Onions pattern. The TWDTW method cannot distinguish the delayed Onions temporal profiles from WinterWheat temporal profiles. This situation led to the confusion of Onions and WinterWheat. This kind of confusion also happened between Onions and Lettuce. In the calculation of TWDTW distance, these two crops had flexible agricultural practices which

led to longer distance between testing points and patterns. Though the Lettuce class' pattern temporal profile had two growing peaks which was different with one growing peak of the Onions class, the TWDTW method did not take the curve' shape into consideration. It was hard for TWDTW to distinguish whether the distance was from the time lag or the real NDVI value difference. These kinds of errors caused the unsatisfactory PAs results for the Onions and Lettuce. The similar things also happened in Belgiu & Csillik's (2018) study. The confusion between Onions, Lettuce and Durum wheat class caused the low PAs and UAs of these classes. The misclassification in Belgiu & Csillik's (2018) study from other classes to fallow class did not happen in this study as a result of the modified validation samples. Another shortcoming of TWDTW method is the TWDTW contains complexed computation. The computation time increased by 126% from the original DTW method. This problem also described in Belgiu & Csillik's (2018) study. One of the objectives of this study is to reduce the computation time of the DTW-based methods. As the problem of computation time stated above, the TWDTW was selected as the DTW-based method for further study.

4.2.2.2. WDDTW and NTDTW

WDDTW used a weight to control the effect of the first derivative in the calculation of the DTW distance of the two sequences. The effect of the first derivative alleviated the problem namely, wrongly matching. For instance, there are two points who have similar NDVI values located in two sequences. One point is in ascending part of a sequence, the other one is in the descending part of the other sequence. The DTW distance of these two points is short due to the fact that the DTW method considers only their NDVI values. These two points will be wrongly matched. Similar to WDDTW, the NTDTW method also takes the shape features into consideration. NTDTW uses mathematical transforms to replace the first derivative of sequences. In this way, two sequences who have similar shapes will get shorter WDDTW distance. For Alfalfa, OtherHay, and Lettuce, the frequency fluctuations of their NDVI values lead to the positive and negative fluctuations of their first derivatives. The positive and negative fluctuations of the first derivatives enhanced the sequences' shape features, as a result, the accuracies improved. The WDDTW was applied on 20 datasets from different application domains for classification and obtained better performance than the original DTW method (Jeong et al., 2011). Górecki & Łuczak (2014) applied the NTDTW on 47 time series datasets for classification. For most of the time series datasets, the NTDTW method achieved higher accuracies than the original DTW method (Górecki & Łuczak, 2014). In our study, the WDDTW and NTDTW were first implemented for the classification of crop temporal profiles. Compared to the original DTW method, the WDDTW and NTDTW had better performance.

Even if WDDTW and NTDTW improved most of the classes' accuracies, the improvement was not significant. The flexible agricultural practices lead to the diversity of the same crop's curves. Onions, Lettuce and WinterWheat were still misclassified. Many samples of OtherHay were classified as Alfalfa class. WDDTW and NTDTW alleviated errors of classification, but it did not solve the problems caused by the flexible agricultural practices.

In case of the computation time, as the authors claimed in their papers, the WDDTW and NTDTW did not increase the computation time complexity (Bagnall et al., 2017; Górecki & Łuczak, 2014; Jeong et al., 2011). The derivative calculation increased the computation time of WDDTW from DTW by 29.4%. The transform calculation increased the computation time of NTDTW from DTW by 28.4%.

4.2.3. Defining the warping path window

In defining the warping path window, we adopted Sakoe's (1978) method to DTW and TWDTW distance calculation. As the author said, The Sakoe-Chiba Band not only reduces the computation time but also improve the accuracy of DTW (Sakoe, 1978). The effect of the warping path window to the DTW distance calculation is similar to the time constraint of the TWDTW. They all limit the matching points between sequences within a certain range. In our study, the window width changing from 5 to 1 meant that a point on one sequence was permitted to match to 5 points to 1 point on the other sequence in both directions. This limit tolerated the flexible agricultural practices and also alleviated the wrongly matching of two crops who had long time lag.

In the experiment of DTW method with warping path window, the overall accuracies were increasing with the decreasing warping path window widths, which verified Sakoe's (1978) conclusion. For TWDTW method, we tried these 5 widths to find an appropriate width which does not have a high impact upon the classification accuracy. We saved the classification results of each pixel and made subtraction between the original TWDTW classification result and each width result. From Table 4, the width from 5 to 4, no pixels changed in the classification results. From 4 to 3, only 53 pixels changed the results from original TWDTW. Even for the width 1, 3507 of 90,000 pixels' results changed. The slightly changed results of TWDTW with different warping window widths were beyond our expectation. Like the different tuning of parameters in TWDTW, the classification results should change with different time constrains (Maus et al., 2016). This phenomenon could be explained. The tuning of the parameters in TWDTW already constrains most of the TWDTW distance calculation within warping path window of width 1. In this way, the classification results will not be affected by the sets of warping path window width larger than 1. In order not to make the warping path window too extreme, the warping path window width was set to 3 for the following classification. For the width 3, it will not affect the classification result of TWDTW and the computation time reduced from 16398 seconds to 7132 seconds by 56.5%.

4.2.4. Decomposition of phenological patterns

In the decomposition of phenological patterns, we tried to extract the crops' growing seasons from the entire temporal profiles. Compared to the original DTW method, the decomposition of phenological patterns applied on DTW not only reduced the computation time but also increased the classification accuracy. Especially, the confusions among Onions, Lettuce and WinterWheat were greatly alleviated. This is due to the fact that the distance calculated beyond the growing season was omitted. With the decomposed phenological pattern the sample who has similar temporal profile with the pattern will be more easily

classified. The decomposition of patterns emphasized the effect of growing seasons. We preserved the entire temporal profiles of the Sugarbeets class because the Alfalfa class had also similar NDVI values during the growing season from October 2016 to August 2017. This means that the non-growing season also contains information for classification. On the basis of the phenological patterns decomposition, we also implemented the defining of warping path window with width 3 in the classification.

In the pattern decomposition of TWDTW with or without warping path window, there was no obvious improvement on the accuracies. The pattern decomposition was mainly used to reduce the computation time. As we shown before, the TWDTW achieved the highest classification accuracy. However, its shortcoming was the long computation time. For the pattern decomposition of TWDTW without warping path window, the computation time decreased from the original TWDTW 16398s to 5291s by 67.7%. For the pattern decomposition of TWDTW with warping path window width 3, the accuracy improved to 89.1%, the computation time was 4029s, reduced by 75.4% from original TWDTW method.

The decomposition of phenological pattern is knowledge-based for now. Except for the Sugarbeets class, the growing season proved to be the most relevant period for achieving good classification results while reducing the computation time. The decomposition of phenological patterns related the work to time series segmentation. Chandola et al., (2010) segmented the temporal profiles of Moderate Resolution Imaging Spectroradiometer (MODIS) NDVI data to derive phenology indices. Conrad et al. (2011) applied temporal segmentation of MODIS time series to improve the crop classification in Central Asian irrigation systems. They proved that the temporal segments of time series data performed better than the discrete time series data (Conrad et al., 2011). Due to time constraints and the complexity of the crop temporal profiles, we failed to propose a mathematical formula or a logical method to decompose the phenological patterns in a fast and accurate way. The temporal segmentation of time series for automatic extracting of phenological patterns would be the future work.

4.3. Conclusions

In this study, a knowledge-based decomposition of phenological patterns method was proposed to reduce the computation time of DTW-based methods without decreasing the accuracy of classification.

DTW, TWDTW, WDDTW, and NTDTW were applied in the study area for crops classification. For TWDTW, WDDTW and NTDTW, the overall accuracies were improved from original DTW by 53.8%, 16.2% and 7.0% respectively. Among these methods, TWDTW achieved the highest accuracy. DTW had the shortest computation time. Reducing the warping path window increased the accuracy of DTW method and decreased both DTW and TWDTW methods' computation time. The knowledge-based decomposition of phenological patterns further reduced the computation time of DTW and TWDTW, whereas the accuracy of pattern decomposition of TWDTW with window width 3 slightly improved.

To obtain a fast classification result without human intervention, the original DTW method with reduced warping path window could be integrated into operational programs dedicated to cropland mapping and monitoring based on satellite image time series. The knowledge-based pattern decomposition applied to

TWDTW with a reduced warping path window would be a good option for getting higher accuracy in the classification of crop temporal profiles.

LIST OF REFERENCES

REFERENCES

- Ali, A. (2009). Comparison of Strengths and Weaknesses of NDVI and Landscape-Ecological Mapping Techniques for Developing an Integrated Land Use Mapping Approach A case study of the Mekong delta , Vietnam Comparison of Strengths and Weaknesses of NDVI and Landscape-Ecolo. *Thesis*.
- Anh, H., Diego, D., Silva, F., Petitjean, F., Forestier, G., Bagnall, A., ... Keogh, E. (2018). Optimizing dynamic time warping 's window width for time series data mining applications. *Data Mining and Knowledge Discovery*, 32(4), 1074–1120. <https://doi.org/10.1007/s10618-018-0565-y>
- Atzberger, C., & Clement. (2013). Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. *Remote Sensing*, 5(2), 949–981. <https://doi.org/10.3390/rs5020949>
- Bagnall, A., Lines, J., Bostrom, A., Large, J., & Keogh, E. (2017). The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. *Data Mining and Knowledge Discovery*, 31(3), 606–660. <https://doi.org/10.1007/s10618-016-0483-9>
- Baumann, M., Ozdogan, M., Richardson, A. D., & Radeloff, V. C. (2017). Phenology from Landsat when data is scarce: Using MODIS and Dynamic Time-Warping to combine multi-year Landsat imagery to derive annual phenology curves. *International Journal of Applied Earth Observation and Geoinformation*, 54, 72–83. <https://doi.org/10.1016/J.JAG.2016.09.005>
- Bégué, A., Arvor, D., Bellon, B., Betbeder, J., de Abelleira, D., Ferraz, R. P. D., ... Verón, S. R. (2018). Remote sensing and cropping practices: A review. *Remote Sensing*, 10(1), 1–32. <https://doi.org/10.3390/rs10010099>
- 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. <https://doi.org/10.1016/j.rse.2017.10.005>
- Berndt, D. J., & Clifford, J. (1994). Using Dynamic Time Warping to Find Patterns in Time Series. In U. M. Fayyad & R. Uthurusamy (Eds.), *KDD Workshop* (pp. 359–370). AAAI Press.
- 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. <https://doi.org/10.1080/10106049.2011.562309>
- Chandola, V., Hui, D., Gu, L., Bhaduri, B., & Vatsavai, R. R. (2010). Using time series segmentation for deriving vegetation phenology indices from MODIS NDVI data. *Proceedings - IEEE International Conference on Data Mining, ICDM*, 202–208. <https://doi.org/10.1109/ICDMW.2010.143>
- Conrad, C., Colditz, R. R., Dech, S., Klein, D., & Vlek, P. L. G. (2011). Temporal segmentation of MODIS time series for improving crop classification in Central Asian irrigation systems. *International Journal of Remote Sensing*, 32(23), 8763–8778. <https://doi.org/10.1080/01431161.2010.550647>
- Das, S. K. (1978). Memory and Time Improvements in a Dynamic, (6), 0–3.
- Deeb, Z. L., Jannetta, P. J., Rosenbaum, A. E., Kerber, C. W., & Drayer, B. P. (1979). Towards 3-d model-based tracking and recognition of human movement: a multi-view approach. *Journal of Computer Assisted Tomography*, 3(6), 774–778. <https://doi.org/10.1097/00004728-197912000-00012>
- E. Keogh, Q. Zhu, B. Hu, Y. Hao, X. Xi, L. Wei, C. A. R. (2011). UCR Time Series Classification Archive. Retrieved January 10, 2019, from http://www.cs.ucr.edu/~eamonn/time_series_data/

- Edenhofer, O., Sokona, Y., Minx, J. C., Farahani, E., Kadner, S., Seyboth, K., ... Zwickel Senior Scientist, T. (2015). *Climate Change 2014*. Retrieved from https://www.ipcc.ch/pdf/assessment-report/ar5/wg3/WGIIIAR5_SPM_TS_Volume.pdf
- El-Gammal, M. I., Ali, R. R., & Samra, A. R. M. (2013). NDVI Threshold Classification for Detecting Vegetation Cover in Damietta Governorate, Egypt. *Journal of American Science*, *10*(3), 2–7. <https://doi.org/10.1088/1361-6560/aad43f>
- ESA. (2018a). MSI Instrument – Sentinel-2 MSI Technical Guide – Sentinel Online. Retrieved February 19, 2019, from <https://earth.esa.int/web/sentinel/technical-guides/sentinel-2-msi/msi-instrument>
- ESA. (2018b). User Guides - Sentinel-2 MSI - Level-2 Processing - Sentinel Online. Retrieved February 19, 2019, from <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/processing-levels/level-2>
- Faloutsos, C., Ranganathan, M., & Manolopoulos, Y. (1994). Fast subsequence matching in time-series databases. *ACM SIGMOD Record*, *23*(2), 419–429. <https://doi.org/10.1145/191843.191925>
- Gollmer, K., & Posten, C. (1995). Detection of distorted pattern using dynamic time warping algorithm and application for supervision of bioprocesses. *On-Line Fault Detection and Supervision in Chemical Process Industries*, *28*(12), 101–106. [https://doi.org/10.1016/S1474-6670\(17\)45408-5](https://doi.org/10.1016/S1474-6670(17)45408-5)
- Gómez, C., White, J. C., & Wulder, M. A. (2016). Optical remotely sensed time series data for land cover classification: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *116*, 55–72. <https://doi.org/10.1016/j.isprsjprs.2016.03.008>
- Górecki, T., & Łuczak, M. (2014). Non-isometric transforms in time series classification using DTW. *Knowledge-Based Systems*, *61*, 98–108. <https://doi.org/10.1016/j.knosys.2014.02.011>
- 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). <https://doi.org/10.3390/rs8010019>
- Guan, X., Liu, G., Huang, C., Id, X. M., & Liu, Q. (2018). An Open-Boundary Locally Weighted Dynamic Time Warping Method for Cropland Mapping. <https://doi.org/10.3390/ijgi7020075>
- Gullo, F., Ponti, G., Tagarelli, A., & Greco, S. (2009). A time series representation model for accurate and fast similarity detection. *Pattern Recognition*, *42*(11), 2998–3014. <https://doi.org/10.1016/j.patcog.2009.03.030>
- Huete, A., Didan, K., Miura, T., Rodriguez, E. P., Gao, X., & Ferreira, L. G. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices, *83*, 195–213. [https://doi.org/10.1016/S0034-4257\(02\)00096-2](https://doi.org/10.1016/S0034-4257(02)00096-2)
- Itakura, F. (1975). Minimum Prediction Residual Principle Applied to Speech Recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *23*(1), 67–72. <https://doi.org/10.1109/TASSP.1975.1162641>
- Jeong, Y. S., Jeong, M. K., & Omitaomu, O. A. (2011). Weighted dynamic time warping for time series classification. *Pattern Recognition*, *44*(9), 2231–2240. <https://doi.org/10.1016/j.patcog.2010.09.022>
- Keogh, E. J., & Pazzani, M. J. (2001). Derivative Dynamic Time Warping. *First SLAM International Conference On Data Mining (SDM'2001)*, 1–11. <https://doi.org/10.1137/1.9781611972719.1>
- Matton, N., Canto, G., Waldner, F., Valero, S., Morin, D., Inglada, J., ... Defourny, P. (2015). An Automated Method for Annual Cropland Mapping along the Season for Various Globally-Distributed Agrosystems Using High Spatial and Temporal Resolution Time Series. *Remote Sensing*, *7*(10), 13208–13232. <https://doi.org/10.3390/rs71013208>
- Maus, V., Camara, G., Cartaxo, R., Sanchez, A., Ramos, F. M., & de Queiroz, G. R. (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*, *9*(8), 3729–3739. <https://doi.org/10.1109/JSTARS.2016.2517118>

- Myers, C., Rabiner, L., & Rosenberg, A. (1980). Performance Tradeoffs in Dynamic Time Warping Algorithms. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(6), 623–635.
<https://doi.org/10.1109/TASSP.1980.1163491>
- Ning Hu, R. B. D. and G. T. (2003). Polyphonic Audio Matching and Alignment for Music Retrieval 5000 Forbes Avenue Pittsburgh , PA 15213-3891 USA. *Audio*, 185–188.
- Petitjean, F., Forestier, G., Webb, G. I., Nicholson, A. E., Chen, Y., & Keogh, E. (2015). Dynamic Time Warping Averaging of Time Series Allows Faster and More Accurate Classification. *Proceedings - IEEE International Conference on Data Mining, ICDM, 2015-Janua*(January), 470–479.
<https://doi.org/10.1109/ICDM.2014.27>
- 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.
<https://doi.org/10.1109/TGRS.2011.2179050>
- Rabi, K. (2017). Planting Period Guide Optimal Acceptable Not Recommended, 2.
- Ratanamahatana, C. A., & Keogh, E. (2004). Making Time-series Classification More Accurate Using Learned Constraints. *Proceedings of the 2004 SLAM International Conference on Data Mining*, 11–22.
<https://doi.org/10.1137/1.9781611972740.2>
- Reed, B. C., Brown, J. F., VanderZee, D., Loveland, T. R., Merchant, J. W., & Ohlen, D. O. (1994). Measuring phenological variability from satellite imagery. *Journal of Vegetation Science*, 5(5), 703–714.
<https://doi.org/10.2307/3235884>
- Roerink, G. J., Menenti, M., & Verhoef, W. (2000). Reconstructing cloudfree NDVI composites using Fourier analysis of time series. *International Journal of Remote Sensing*, 21(9), 1911–1917.
<https://doi.org/10.1080/014311600209814>
- Sakoe, H. (1978). *Dynamic Programming Algorithm Optimization for Spoken Word Recognition*. *IEEE TRANSACTIONS ON ACOUSTICS, SPEECH, AND SIGNAL PROCESSING*. Retrieved from https://pdfs.semanticscholar.org/18f3/55d7ef4aa9f82bf5c00f84e46714efa5fd77.pdf?_ga=2.170604520.955786915.1534147014-387209061.1534147014
- Sakurai, Y., Faloutsos, C., & Yamamuro, M. (2007). Stream monitoring under the time warping distance. *Proceedings - International Conference on Data Engineering*, 1046–1055.
<https://doi.org/10.1109/ICDE.2007.368963>
- Silva, D. F., & Batista, G. E. A. P. A. (2016). Speeding Up All-Pairwise Dynamic Time Warping Matrix Calculation. *Proceedings of the 2016 SLAM International Conference on Data Mining*, 837–845.
<https://doi.org/10.1137/1.9781611974348.94>
- Tan, C. W., Herrmann, M., Forestier, G., Webb, G. I., & Petitjean, F. (2018). Efficient search of the best warping window for Dynamic Time Warping. *Proceedings of the 2018 SLAM International Conference on Data Mining*, 225–233. <https://doi.org/10.1137/1.9781611975321.26>
- Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8(2), 127–150. [https://doi.org/10.1016/0034-4257\(79\)90013-0](https://doi.org/10.1016/0034-4257(79)90013-0)
- University of California Cooperative Extension. (2017). vegetable crops planting & harvesting calendar. Retrieved from <http://ceimperial.ucanr.edu/files/206960.pdf>
- USDA. (2010). Field Crops Usual Planting and Harvesting Dates. *Agricultural Handbook*, (628), 41.
<https://doi.org/October 2010>
- USDA. (2017a). California Agricultural Statistics. *Annual Bulletin*. Retrieved from http://www.nass.usda.gov/Statistics_by_State/California/Publications/California_Ag_Statistics/2013cas-all.pdf
- USDA. (2017b). USDA - National Agricultural Statistics Service - Research and Science - CropScape and Cropland Data Layers. Retrieved January 21, 2019, from https://www.nass.usda.gov/Research_and_Science/Cropland/sarsfaqs2.php#Section3_22.0

- USDA. (2018). CropScape - NASS CDL Program. Retrieved February 20, 2019, from <https://nassgeodata.gmu.edu/CropScape/>
- Waldner, F., Canto, G. S., & Defourny, P. (2015). Automated annual cropland mapping using knowledge-based temporal features. *ISPRS Journal of Photogrammetry and Remote Sensing*, 110, 1–13. <https://doi.org/10.1016/J.ISPRSJPRS.2015.09.013>
- Yan, L., & Roy, D. P. (2014). Automated crop field extraction from multi-temporal Web Enabled Landsat Data. *Remote Sensing of Environment*, 144, 42–64. <https://doi.org/10.1016/j.rse.2014.01.006>
- Zhu, Y., Shasha, D. E., & Zhao, X. (2003). Query by Humming - in Action with its Technology Revealed. *SIGMOD Conference*, 675. Retrieved from <http://dblp.uni-trier.de/db/conf/sigmod/sigmod2003.html#ZhuSZ03>