AIMS Geosciences, 2 (4): 302–328
DOI: 10.3934/geosci.2016.4.302
Received date 31 July 2016
Accepted date 24 October 2016
Published date 09 November 2016

http://www.aimspress.com/journal/geosciences

Research article

Object-based Dimensionality Reduction in Land Surface Phenology Classification

Brian E. Bunker [1], Jason A. Tullis [2,*], Jackson D. Cothren [2], Jesse Casana [3], and Mohamed H. Aly [2]

[1] Environmental Systems Research Institute (Esri), 3060 Little Hills Expressway, St. Charles, MO, USA
[2] Department of Geosciences and Center for Advanced Spatial Technologies (CAST), University of Arkansas, Fayetteville, AR, USA
[3] Dartmouth College, Hanover, NH, USA

* **Correspondence**: Email: jatullis@uark.edu; Tel: 479-575-8784.

Abstract: Unsupervised classification or clustering of multi-decadal land surface phenology provides a spatio-temporal synopsis of natural and agricultural vegetation response to environmental variability and anthropogenic activities. Notwithstanding the detailed temporal information available in calibrated bi-monthly normalized difference vegetation index (NDVI) and comparable time series, typical pre-classification workflows average a pixel's bi-monthly index within the larger multi-decadal time series. While this process is one practical way to reduce the dimensionality of time series with many hundreds of image epochs, it effectively dampens temporal variation from both intra and inter-annual observations related to land surface phenology. Through a novel application of object-based segmentation aimed at spatial (not temporal) dimensionality reduction, all 294 image epochs from a Moderate Resolution Imaging Spectroradiometer (MODIS) bi-monthly NDVI time series covering the northern Fertile Crescent were retained (in homogenous landscape units) as unsupervised classification inputs. Given the inherent challenges of *in situ* or manual image interpretation of land surface phenology classes, a cluster validation approach based on transformed divergence enabled comparison between traditional and novel techniques. Improved intra-annual contrast was clearly manifest in rain-fed agriculture and inter-annual trajectories showed increased cluster cohesion, reducing the overall number of classes identified in the Fertile Crescent study area from 24 to 10. Given careful

segmentation parameters, this spatial dimensionality reduction technique augments the value of unsupervised learning to generate homogeneous land surface phenology units. By combining recent scalable computational approaches to image segmentation, future work can pursue new global land surface phenology products based on the high temporal resolution signatures of vegetation index time series.

Keywords: object-based segmentation; time series; unsupervised clustering; land surface phenology; MODIS

1. Introduction

Land surface phenology describes seasonal changes in vegetation, from a remote sensing perspective, at regional and global scales [1,2]. From ecosystem and biodiversity to climate change applications, large scale vegetation phenological trends have been linked to predicted onset of drought in arid climates [3,4], and to observed changes in growing season length over multiple decades [1]. The northern Fertile Crescent, a region of historical and modern significance covering portions of Syria, Lebanon, Turkey, and Iraq, clearly benefits from an increased understanding of vegetation phenology due to persistent reliance on scarce water resources [5,6].

Regional agricultural boundaries (e.g., zones suitable for dry farming of staple cereals such as wheat and barley in the northern Fertile Crescent), are often based on generalized precipitation regimes and limited vegetation phenology observed *in situ* [7]. However, remote sensing-derived land surface phenology zones offer broad coverage and detail, and by systematic observation, account for both environment and anthropogenic activities [8,9]. This study complements past work to improve classification of land cover through satellite image time series [10,11] by offering a new perspective on incorporating underused multi-decadal seasonal temporal data while creating phenological zones or homogeneous landscape units [9]. Toward this goal, we describe an object-based spatial dimensionality reduction approach when clustering MODIS NDVI time series. Our novel method is compared to the commonly used "mean year" dimensionality reduction technique [9,12,13] to highlight improved preservation of seasonality and multi-year trajectories within phenological zones. The new method, while demonstrated in the northern Fertile Crescent, may be extended to other regions or globally to augment the value of unsupervised classification in current and future land surface phenology applications.

1.1. Time Series Imagery

Satellite remote sensing has produced records of Earth's land conditions for multiple decades, made possible by regular revisit schedules inherent in repeating processional orbits. Notable examples of continuing programs that have lasted longer than a decade include Landsat, jointly sponsored by U.S. Geological Survey (USGS) and National Aeronautics and Space Administration (NASA), Advanced Very High Resolution Radiometer (AVHRR) sponsored by National Oceanic and Atmospheric Administration (NOAA), and NASA's Moderate Resolution Imaging Spectroradiometer (MODIS).

Unlike Landsat with its 16-day revisit period and 30 m spatial resolution, the latter two sensors exhibit higher temporal resolutions and support land surface phenology applications with inherently coarser spatial resolution requirements. The AVHRR sensor program, initiated in 1978, has acquired twice-daily global imagery at a spatial resolution of 1.1 km. MODIS was first placed in orbit in 1999 with the capability of daily image collection at a maximum spatial resolution of 250 m for its red and near-infrared bands. The complete archive of successive temporal images acquired by each mission is comprised of up to multi-decadal records that have garnered significant interest in both academia and industry [14]. The NOAA Suomi National Polar-orbiting Operational Environmental Satellite System (NPOESS) Preparatory Project (NPP), launched in 2011 with additional launches planned, includes the Visible Infrared Imaging Radiometer Suite (VIIRS) with many similarities to MODIS.

Agency pre-processing procedures applied to the archive of remotely sensed images have created internally consistent time series which are ready-to-use in subsequent research and monitoring applications. Advances in image correction (e.g., for atmospheric attenuation, satellite drift, bi-directional reflectance, etc.) have greatly benefited time series imagery analysis. Holben [15] presented Maximum Value Compositing (MVC), which is applied to time series imagery by retaining a maximum reflectance value for each pixel over a multi-day compositing period. Common compositing periods range from one week to one month in order to ensure acquisition of maximum reflectance for each pixel in a scene, and may only represent certain seasons of the year [13]. The compositing periods, while long enough to gain valuable cloud-free per-pixel data, preserve seasonal vegetation reflectance curves necessary for phenological studies. Early work on AVHRR image data to calibrate sensor equipment and correct atmospheric interference produced time series with greatly reduced sensor and environmental pixel degradation [16–18]. Pinzón et al. [19] developed a method using empirical mode decomposition to isolate and remove spectral artifacts in time series imagery caused by gradually changing solar zenith angles due to satellite platform orbital drift. These and other techniques have been developed to create consistent time series imagery such as AVHRR, MODIS, and Satellites Pour l'Observation de la Terre (SPOT) Vegetation (SPOT-VGT) that has been so valuable in supporting land surface phenology research in recent decades [1].

1.2. Vegetation indices in time series

The commonly utilized NDVI [22,23] takes advantage of high chlorophyll absorption in the red and foliar reflectance in near infrared wavelengths of healthy green vegetation. During the more than three decades from 1978-present, daily NDVI (with values from -1 to 1) images are easily calculated using near-infrared and red bands from AVHRR and/or MODIS time series (Figure 1):

$$NDVI = \frac{\rho_{NIR} - \rho_{red}}{\rho_{NIR} + \rho_{red}} \quad (1)$$

where ρ_{red} is Band 1 reflectance (0.580–0.680 μm for AVHRR and 0.620–0.670 μm for MODIS), and ρ_{NIR} is Band 2 reflectance (0.725–1.10 μm for AVHRR and 0.841–0.876 μm for MODIS). Despite the refined spectral resolution of the red and near-infrared MODIS bands, a calibrated NDVI

time series for AVHRR that is compatible with MODIS-derived NDVI time series is available [21]. Furthermore, when compared with the very similar Simple Ratio which is dependent upon the same spectral bands, NDVI itself is well suited to the lower biomass conditions [14] of the dry farming transition zone between the Middle East steppe and the Mesopotamian shrub as identified in the 2005 World Wildlife Fund Terrestrial Ecoregions map.

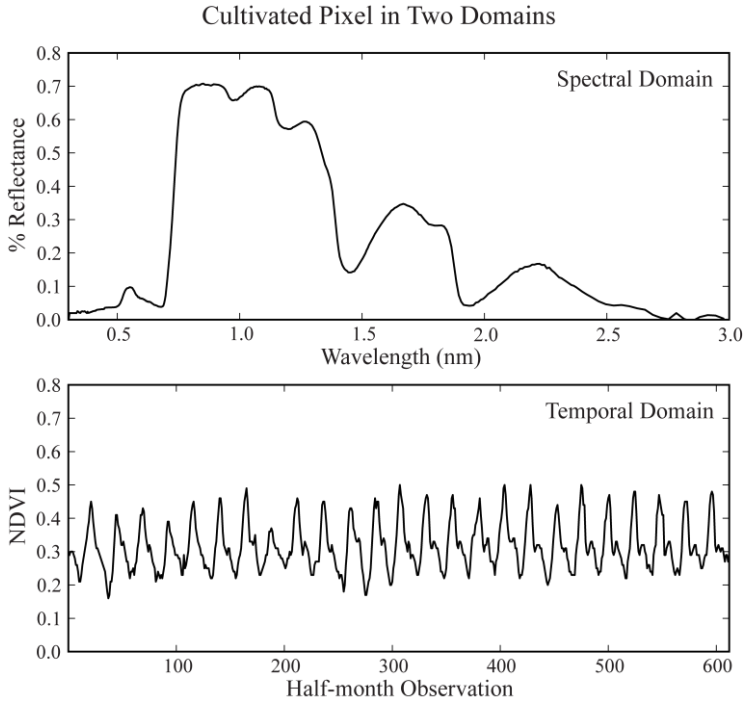


Figure 1. Possible values for a cultivated pixel in both spectral (top) and temporal (bottom) domains. Spectral curve data (top) was obtained from the USGS Digital Spectral Library [20]. Each half-month observation (bottom) is taken from a 25-year AVHRR NDVI time series produced by Tucker et al. [21]. Each pixel may contain unique reflectance characteristics in both spectral and temporal domains, allowing for classification by spectral or phenological similarity.

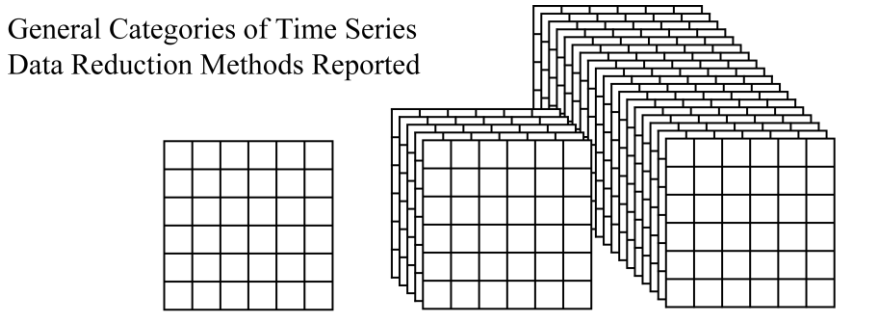
The natural phenological cycle, observable to humans through the repeating seasonal changes of vegetation, produces a seasonal oscillation (often referred to as a signal, curve, or profile) in NDVI calculated for the same pixel over consecutive years (Figure 1). Time series NDVI data spanning the length of a growing season or longer can serve as a valuable proxy for plant phenology [24].

1.3. Time Series Vegetation Index Classification

Time series vegetation indices such as NDVI data are used as input for unsupervised classification algorithms to create relatively homogeneous phenological data partitions or landscape units [9,12]. Since each input pixel corresponds to one spatial location and many temporal observations, input to the classification algorithm is a vector (array) of data in *n* dimensions where *n* is the number of images in the time series. Each NDVI vector plotted by time reveals the phenological curve at each pixel location (Figure 1). Using a measure of statistical separability, a preferred classification algorithm groups and

assigns class membership labels to pixel vectors of similar phenological curve shape. However, given the need for dimensionality reduction, applications of time series imagery or indices rarely incorporate full temporal curve shape into the classification product.

Researchers have approached classification of satellite imagery-derived time series using several methods. Common approaches are divided into three categories based on the form of the time series to be incorporated as input for classification: flattened, reduced, and full-data time series (Figure 2). First, flattened-data time series are characterized by the replacement of original data time series with a single raster of representative values or metrics corresponding to important phenological events. One such metric is "length of growing season" calculated as the number of days with an NDVI value greater than a predetermined threshold [8]. Second, reduced-data time series are designed to reduce the dimensionality of a dataset while maintaining a multi-raster temporal sequence. A common reduction method is to calculate mean NDVI for each month represented by a multi-year time series [9,12,13]. Finally, full-data time series utilize the maximum number of available imagery dates to obtain temporal NDVI vectors (arrays of values) for every pixel location. Only full-data time series precisely track the trajectory of observed data values over time [25], adding valuable insight to anthropogenic and environmental spatial variability.



	Flattened	*Reduced*	*Full*
Description	Full time series is used to calculate a representative value or metric for each pixel	Full time series is used to create new rasters of reduced dimensionality in either temporal or spatial domain	Minimal or no data reduction is applied to produce a regularly spaced time series between start date and end date
Result	Single raster	Raster set with temporal sequence	Raster set with temporal sequence
Examples	Average start date or end date of season, date of max or min NDVI, max or min NDVI value, etc.	Mean NDVI per repeating temporal frame (e.g. mean Jan, Feb, etc. value for each year in original time series)	GIMMS AVHRR NDVI time series or MODIS MOD13Q1 NDVI time series

Figure 2. In general, time series of remotely sensed images and derived indices have been utilized as full-, reduced-, and flattened-data time series in various application workflows. Each approach to time series data reduction may target specific attributes of land surface vegetation phenology.

Multiple researchers have encountered the need to reduce a vegetation index time series as part of a classification or clustering workflow. For example, Kouchoukos [12] produced an "agro-ecological" map from fifteen years of time series data in Mesopotamia. He calculated monthly averages of AVHRR NDVI data over all fifteen years, resulting in a reduced set of twelve NDVI images that represent an average year. Kouchoukos then temporally ordered and subjected the monthly averages to unsupervised

classification, resulting in a map with seven relatively homogeneous land surface phenology classes. In a second example, Al-Bakri & Taylor [26] repeated this method of data reduction by averaging corresponding months over time. Other techniques employed by scholars include principal component analysis (PCA) [9,27] and Fourier transform [24,28] to reduce dimensionality and to remove noise while maintaining a partial temporal sequence relative to the original time series data.

Object-based image analysis (OBIA) and image segmentation (e.g., for primary object detection through Trimble's eCognition) continue to be most often associated with relatively high spatial resolution applications, such as automated landslide detection [29]. However, a few recent studies have utilized object-based image segmentation in conjunction with remote sensing-derived vegetation index time series. For example, Bontemps et al. [30] demonstrated the use of multi-temporal image segments derived from a SPOT-VGT time series to create change/no change objects in order to better take into account temporal dependencies often ignored in traditional change detection algorithms. As part of a land cover study in Namibia, Hüttich et al. [31] segmented Landsat ETM+ imagery as a scaling mechanism to link *in situ* reference data with relatively coarse spatial resolution MODIS imagery. In an investigation of wildfire susceptibility in Sardinia, Italy, De Angelis et al. [32] first applied the mean year method to reduce the dimensionality of a MODIS 250 m NDVI time series (specifically the MOD13Q1 product) with 253 epochs down to 23 mean NDVI images. This intermediate product was then subjected to multiresolution image segmentation in eCognition to produce homogenous landscape units suitable for extraction of land surface phenological metrics and subsequent fire ignition-related cluster analysis. Zhong et al. [33] extracted and analyzed land surface phenology metrics within segments pre-computed from finer spatial resolution data. In a final example, Bisquert et al. [9] extracted image segments from a PCA-reduced selection of indices and dates from multiple vegetation and Haralick texture time series. They concluded that resulting homogenous landscape units produced are of significant value in a variety of vegetation and ecosystem contexts.

1.4. Statement of the Problem

Given the advent of object-based methods in high temporal resolution remote sensing workflows, a key question is how best to reduce the time series prior to information extraction. This work presents a novel time series data reduction method and compares it to the commonly used mean year reduction method as input to land surface phenology classification. Our approach maintains temporal fidelity after an object-based image segmentation process is applied. This study focuses on the information content added to subsequent unsupervised clusters when long time series are utilized. With the northern Fertile Crescent and its traditional dry farming and scarce water resources as context, three specific objectives were successfully pursued to 1) characterize differences between unsupervised clusters or classes based on mean year reduced versus segment mean reduced long time series NDVI; 2) identify differences in the spatial distribution of classes; and 3) determine how varying class boundaries produced using the two data reduction methods contribute to our understanding of regional and temporal phenological patterns within the northern Fertile Crescent.

2. Materials and Methods

2.1. Unsupervised Classification

Unsupervised classification (or clustering) groups data points into classes according to the inherent structure of the data in measurement space and is usually based on an algorithm with a user-defined *k* parameter that specifies how many classes are anticipated. The expert who operates the relevant classification software has little to no additional input. The classification software typically utilizes a series of heuristic evaluations of intermediate data class membership to adjust class parameters and reassign data to new classes while the algorithm converges on the best distribution of data given *k* target classes. Traditionally, *k* has been derived from a previous classification of the area of interest or interactively adjusted until the analyst determines that useful classes have been identified. In contrast, relatively few studies using unsupervised classification of NDVI incorporate *a posteriori* class similarity statistical analysis described by Tou and Gonzales [34] and Swain and Davis [35], and exemplified by Nguyen et al. [25].

Cluster validation is applied *a posteriori* as a method for evaluating the effectiveness of unsupervised classification, but separate and distinct from final accuracy assessment using Cohen's kappa coefficient [36] or other analysis of the classification error matrix [37]. The goal of cluster validation is to answer the question: did the classification algorithm identify the structure and number of inherent clusters, or classes, in the input data? Methods to answer this question utilize indices of a) class cohesion and b) class separation. Class cohesion describes the compactness of the class members around their class centers, while class separation measures the uniqueness of each class or distance between classes. Researchers have developed many cluster validation indices (CVIs) to manage various scenarios [38].

Nguyen et al. [25] apply the *divergence index* defined by Swain and Davis [35]. The index uses class signature files to calculate separability between classes. Signature files contain a mean vector and covariance matrix for every class in a classification scheme. These values are calculated using class assignments and the original data.

A transformed divergence index [35,39] is calculated as:

$$TD_{avg} = \Sigma_{i=1}^{m}\Sigma_{j=1}^{m} p(\omega_i)\, p(\omega_j) TD_{ij} \tag{2}$$

where

$$TD_{ij} = 2000\left(1 - \exp\left(\frac{-D_{ij}}{8}\right)\right) \tag{3}$$

and

$$D_{ij} = \frac{1}{2} tr\left[\left(\Sigma_i - \Sigma_j\right)\left(\Sigma_j^{-1} - \Sigma_i^{-1}\right)\right] + \frac{1}{2} tr\left[\left(\Sigma_i^{-1} + \Sigma_j^{-1}\right)\left(U_i - U_j\right)\left(U_i - U_j\right)^T\right] \tag{4}$$

Elements of equations 2, 3, and 4 are:

1. TD_{ij}—transformed divergence between classes i and j
2. $p(\omega_i)$—*a priori* probability of class membership; equal to m^{-1}

3. m —total number of classes
4. D_{ij} —the divergence between classes i and j
5. $tr[x]$ – the trace of matrix x; the sum of the elements on the diagonal of x
6. Σ_i —the covariance matrix of class
7. U_i —the mean vector of class
8. T —the transpose function

Minimum transformed divergence is defined as the minimum transformed divergence measurement between a pair of classes in the set of all class pairs as calculated in Equation 3. It is important to note that the minimum transformed divergence serves as a complimentary measure to the average transformed divergence when evaluating classification schemes. The minimum transformed divergence measures how well each classification scheme (or *a priori* number of classes specified) separates the most closely related classes. A relatively low value indicates that the closest (most similar) classes are not well separated, while a relatively high value gives the analyst more confidence in the classification scheme. In general, however, the more classes the data are divided into, the lower the minimum transformed divergence. The analyst must therefore balance the increasing average transformed divergence against the decreasing minimum transformed divergence to select an optimal classification scheme.

2.2. Image Segmentation

The value derived from hierarchically grouping data into mutually exclusive subsets has long been recognized [40]. Increasingly popular image segmentation is the process of decomposing imagery into homogenous regions called segments that can be arranged as spatially hierarchical objects. This mimics the cognitive ability of the human visual processing system to identify objects by recognizing homogeneity based on proximal data values of similar magnitude [41]. Multi-resolution image segmentation introduced by Baatz & Schäpe [41] maximizes homogeneity within segments *and* maximizes heterogeneity between segments. Homogeneity h for image segments is defined as:

$$h = \sqrt{\Sigma_d \left(\frac{f_{1d} - f_{2d}}{\sigma_{fd}} \right)} \qquad (5)$$

where f_{1d} is the feature value for segment 1 in dimension d, and σ is the standard deviation of feature f for all segments in dimension d. The feature value f is a metric that can be derived from the image object. Subscript d refers to the image dimension (e.g., band, or in the present context, date in a time series). For example, in a time series NDVI context, f could be the mean (or variance,

etc.) of all the March 15th, 2016 NDVI pixel values in a segment. Homogeneity would be calculated across all dimensions of the image stack; in a calibrated time-series context, the image stack is comprised of a series of sequential time frames (e.g., 15 February, 1 March, 15 March, etc. in a specific year).

To minimize the heterogeneity within segments, multiresolution segmentation merges pairs of image objects with minimal change of heterogeneity, h_{diff}, among possible merges (Figure 3). This is measured by the equation:

$$h_{diff} = \Sigma_d w_d (n_1(h_{md} - h_{1d}) + n_2(h_{md} - h_{2d})) \quad (6)$$

where h_{1d} is the homogeneity of segment 1 in dimension d, h_{md} is the homogeneity of segments 1 and 2 after a virtual merge, n_1 is the size of object 1, and w_d is the weight of dimension d. The virtual merge is a "*what if*" condition: *what* will the homogeneity be *if* segment 1 and 2 are merged? The "*what if*" scenario is repeated to calculate h_{diff} for all pairs of adjacent segments, and the minimum value is chosen to define the actual segment merge.

Because image segmentation groups homogenous values, it effectively reduces the dimensionality of data needed to represent an image. While not easily compared with spectral dimensionality reduction approaches such as band or feature selection [42], segmentation is conceptually similar to resampling of raster data to a larger cell size. However, instead of performing an arbitrarily uniform gridded merging of cells, image segmentation merges cells according to value homogeneity. The mean value of the merged pixels can be used to represent the new, larger area segment. Reduced raster data appear as *s* segments instead of *p* pixels, where *s* is always less than *p*.

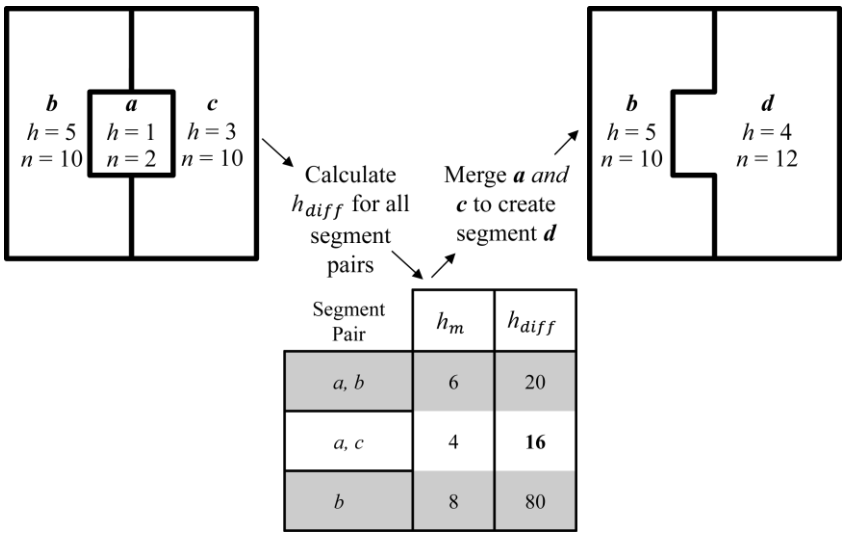


Figure 3. Imagery including satellite-derived time series can be segmented to create image objects bounding neighboring homogeneous pixels. This example illustrates the segment merge decision criteria: minimize potential increase in heterogeneity within a segment.

2.3. Study Area

This study focused on a region of the Middle East referred to as the northern Fertile Crescent, covering the area extending longitudinally from 28˚ to 50˚ east with a latitudinal span from 28˚ to 42˚ north. The Fertile Crescent, a term coined by archaeologist James Henry Breasted in 1906, describes an arc-shaped, agriculturally productive zone trending east-west with a central northern apex and southward-bowed ends. The region is characterized by a Mediterranean climate, with hot, dry summers and cool, wet winters with as much as 90% of annual precipitation falling between November and March. Alongside this seasonality, a principle feature of the region is strong precipitation gradients, with annual totals along the humid Mediterranean littoral and in the northern Taurus-Zagros Mountain range topping 1,800 mm, rapidly decreasing to less than 100 mm in the central Syrian Desert less than 300 km away [43]. This pattern of rainfall, driven largely by orographic effects on Mediterranean and North Atlantic climate systems, creates a similarly strong natural vegetation gradient [5] and has profound effects on the spatial distribution and sustainability of agriculture [7,44]. While the anastomosing Tigris-Euphrates river system in southern Mesopotamia supports widespread irrigation agriculture, the broad plains of the northern Fertile Crescent feature deeply incised rivers that result in very limited opportunities for irrigation (Figure 4). In this area, most agriculture has traditionally been dependent on rainfall, and thus the spatial patterning of precipitation and other water resources, particularly in more arid areas, is a key factor in determining agricultural potential, both ancient and modern.

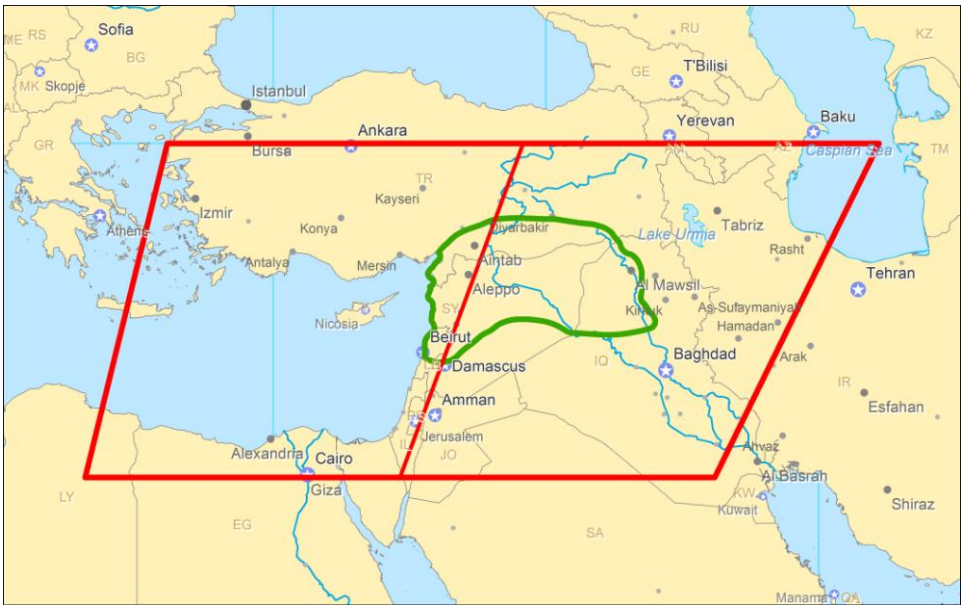


Figure 4. The northern Fertile Crescent study area outlined in green, with the extents of the two adjoining sinusoidal MODIS tiles in red. Basemap courtesy of DeLorme and Esri.

As with studies of modern agriculture, archaeologists have similarly utilized rather static models of agricultural sustainability, based largely on the location of the 250 mm annual rainfall isohyet, below which it is believed dry-farming of wheat and barley is not possible. Ancient settlements

located below the modern boundary of rain fed farming have been taken as evidence of past climate change, while the survival or collapse of settlements located above the 250 mm isohyet have similarly been linked to variability in annual rainfall [45,46]. As a region that is home to the world's first sedentary agricultural communities more than 10,000 years ago [47] as well as to some of the first cities [48], a better understanding of landscape phenology in the northern Fertile Crescent offers a key contribution to longstanding debates regarding the origin and development of early complex societies. A refined picture of spatial and temporal patterns in water availability and vegetation health similarly provides a basis for more nuanced analysis of modern agricultural production in this extremely politically volatile region.

2.4. Remotely Sensed Data

2.4.1. MODIS

MODIS data within two adjoining sinusoidal grid tiles were obtained for twelve years (2000–2012) from the USGS Land Processes Distributed Active Archive Center (LP DAAC) using NASA's [49] Reverb web interface. Individual scenes of the 250 m MODIS vegetation index time series, MOD13Q1, were examined for cloud and/or atmospheric contamination and poor sun-target-sensor geometry. Pixels without contamination or poor geometry were retained for the vegetation index time series based on 16 day MVCs. Solano et al. [50] explain the full processing procedure of the MOD13Q1 data.

The downloadable scenes were multi-layer and conveyed the actual data values for NDVI and Enhanced Vegetation Index (EVI) along with pixel quality. NDVI was stored as 16-bit integers with a range of -2,000 to 10,000. Actual NDVI were calculated from the stored values using the following equation:

$$NDVI = raw \times 0.0001 \quad (7)$$

"NoData" values were presented as -3,000, while data quality values were stored as 8-bit integers ranging from -1 to 3 (Table 1). Data values with corresponding error values of 0 through 2 were used in this study.

Table 1. Quality flag values for MOD13Q1; adapted from Solano et al. [50].

Value	Summary	Description
-1	Fill/no data	Not processed
0	Good data	Use with confidence
1	Marginal data	Useful, but look at other quality information
2	Snow/ice	Target covered with snow/ice
3	Cloudy	Target not visible, covered with clouds

For each acquisition date, the two MODIS tiles were mosaicked using the Esri's ArcPy Python package. The complete MODIS time series was comprised of 294 mosaicked NDVI rasters corresponding to 294 compositing periods over the given 12 years. MODIS data were extracted from

each associated downloaded HDF file and left in their scaled NDVI integer form to accommodate use of 16-bit rasters. This provided a 50% savings in memory storage (from 32 to 16 bits).

2.4.2. Landsat and Other Higher Spatial Resolution Imagery

Selected Landsat Thematic Mapper (TM) scenes, acquired 6 July 1984, 30 August 1984, and 1 August 1985, were also accessed through NASA's [49] Reverb web system. More recent Landsat 8 Operational Land Imager (OLI) imagery acquired 24 Apr 2013 was also accessed using USGS's [51] Global Visualization Viewer (GLOVIS) tool. Additionally, recent high spatial resolution aerial and satellite imagery, available through Esri's ArcMap, were accessed to aid manual interpretation of classification results.

2.5. Data Reduction

2.5.1. Mean Year Method

Reduced-data mean year time series (Figure 5) were created from the MODIS data using ArcPy. Pixels in scenes that had undesirable quality flag values were given a value of "NoData" so as to not influence the calculated means. Corresponding time frames for each year in the time series were averaged to produce a "mean year" time series. The mean year consisted of 23 rasters because the compositing time of the MODIS data is 16 days.

The process used to calculate the mean year time series was as follows:

1. Identify compositing periods for each year (i.e. Jan 1–16, Jan 17–Feb 1… etc.).
2. Group all images in the time series by their compositing period.
3. Calculate local sum of available NDVI values within each compositing period.
4. Calculate local sum of the number of valid NDVI values within each compositing period.
5. Calculate local division of step 3 result by step 4 result.
6. Repeat steps 3–5 for each compositing period in the year.

The resulting set of mean compositing period rasters were temporally ordered as mean year NDVI time series (23 layers).

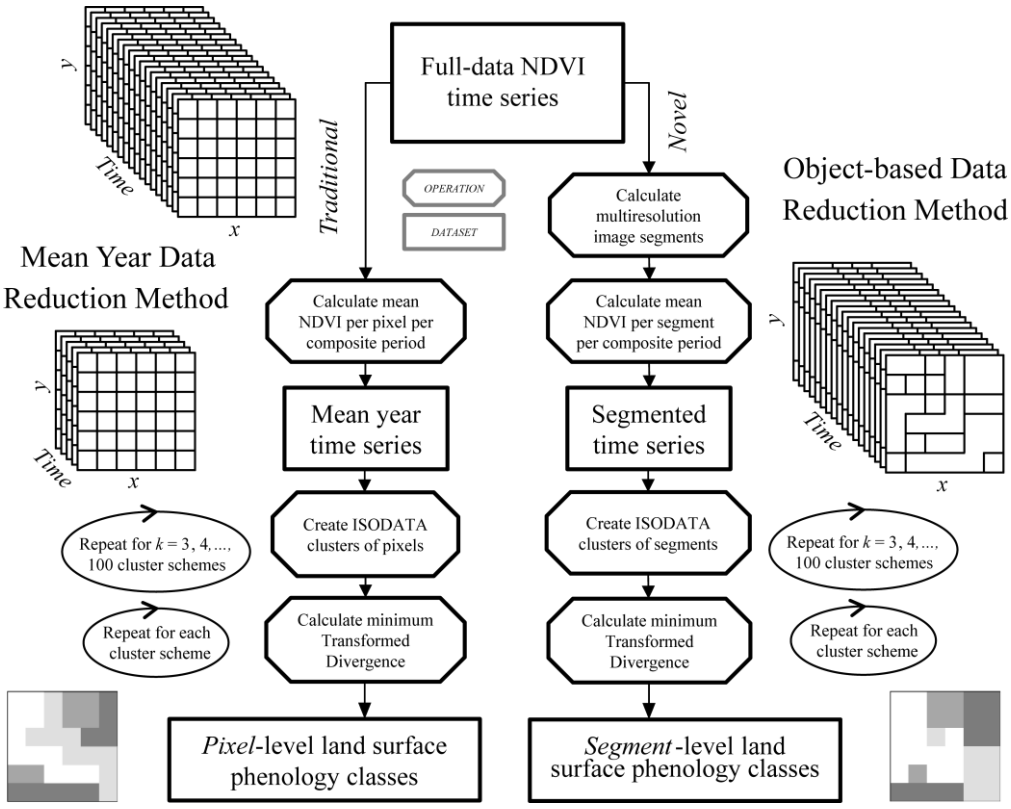


Figure 5. Generalized workflow comparison between the traditional mean year data reduction method (left branch) and the novel object-based data reduction method (right branch). The fundamental difference is the initial multiresolution image segmentation in the novel method, allowing for spatial as opposed to temporal aggregation. Each method was tested with NDVI time series from MODIS.

2.5.2. Object-based Method

Trimble's [52] eCognition was used to create image segments (objects) from the MODIS time series as a novel reduction of NDVI time series data (Figure 5). After the MODIS time series images were temporally ordered and loaded into eCognition, each image was split into 127 tiles, each 600 × 600 pixels, due to the large amount of calculations and memory required for segmentation of the full time series. Multi-resolution segmentation was executed for each MODIS time series tile using image pixels as input. Shape and compactness parameters were set to zero which allowed for both compact and linear features (e.g., the Tigris River Valley) to be segmented. Scale parameter options were heuristically tested as explained below. The MODIS tiles were subsequently stitched together using rules that identified tile border objects, and re-segmented new objects in these areas using the original pixel data and specified object parameters (scale, shape, and compactness). This was done to remove superimposed linear segments introduced in the tiling phase (Figure 6).

Multiresolution segmentation was repeated for incremental scale parameters of 15, 18, 21, 24, 27, 30, 33, 36, 39, and 42 for MODIS data. The segments produced with each scale parameter were compared to the total number of pixels in a single date of imagery to determine the percent of data that was reduced and the average number of pixels per segment that were created under each proposed

segmentation scheme. Manual selection of the best scale parameters versus amount of data reduction was performed using graph comparisons (Figure 8). Maximizing percent data reduction was paramount in the selection criteria for memory-intensive processing of a MODIS time series. (This numerical analysis could be automated to reduce time cost barriers while employing this method.) The selected segments were then exported from eCognition to Esri's Shapefile format.

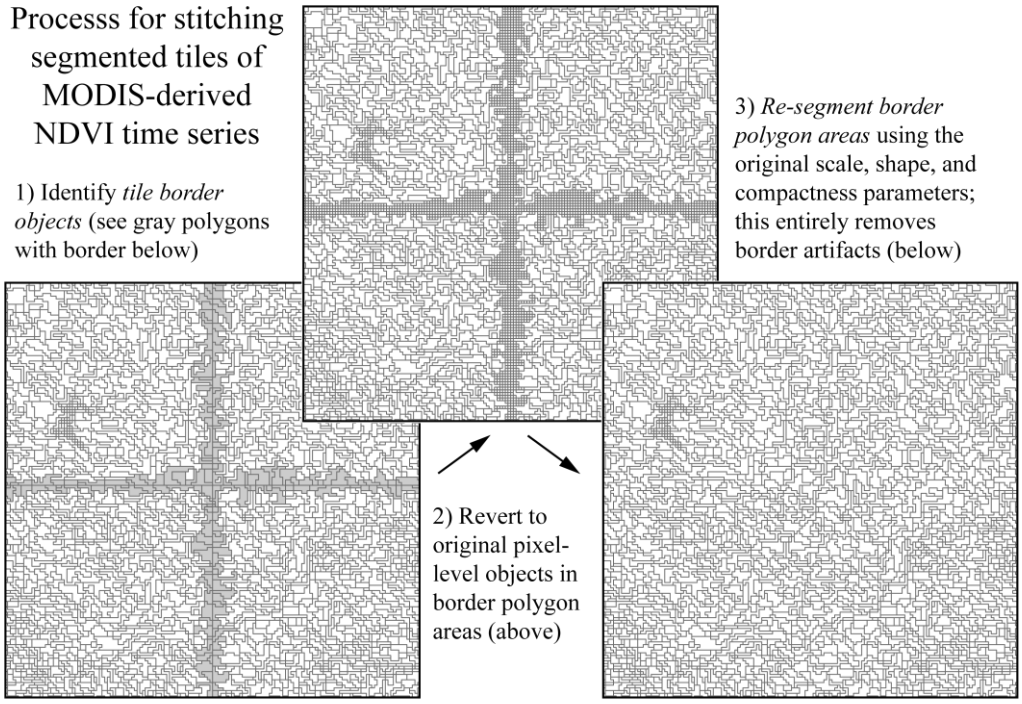


Figure 6. Segmented tiles of MODIS-derived NDVI time series were further processed to remove linear segment boundary artifacts introduced in the tiling process (required for computational scaling in large area applications).

To prepare the data for subsequent unsupervised classification in ERDAS Imagine, the exported segment polygons were used to create a series of spatially un-registered rasters (one for each date in the time series). Each raster cell contained the mean value of the pixels bounded by the segment and calculated from the corresponding NDVI image in the time series. The resulting time series product contained a raster for every compositing period of the full time series, but whose data were aggregated by segment.

2.6. Unsupervised Classification (Clustering)

Each pixel location in the mean year time series represented an NDVI vector spanning a single "mean" year. Layers were arranged so that the beginning of the calendar year was layer 1 and subsequent time frames were assigned ascending layer numbers for a total of 23. In contrast, each pixel location in the segment mean time series represented the temporal trajectory vector of NDVI spanning 294 compositing periods. Layers in the data set were temporally arranged so the earliest

compositing time and year were assigned to the first layer and subsequent time frames were assigned to ascending layers. Both mean year reduced-data and segment mean reduced-data time series were subjected to unsupervised classification using the ISODATA algorithm as implemented by ERDAS Imagine [39].

2.6.1. K parameter Adjustment and Cluster Validity

Thematic accuracy assessment of remote sensing derivatives typically involves sampled reference data collection, the creation of an error matrix, and the calculation of such metrics as overall, producer's and user's accuracies [53]. Unfortunately, comparison and validation of land surface phenology classification methods is challenged by the fact that derived classes are based on remote sensing time series (in the case of this study, 294 compositing periods over 12 years). With no universal reference for population of an error matrix associated with derived land surface phenology classes, this study focused on enabling methodological comparison.

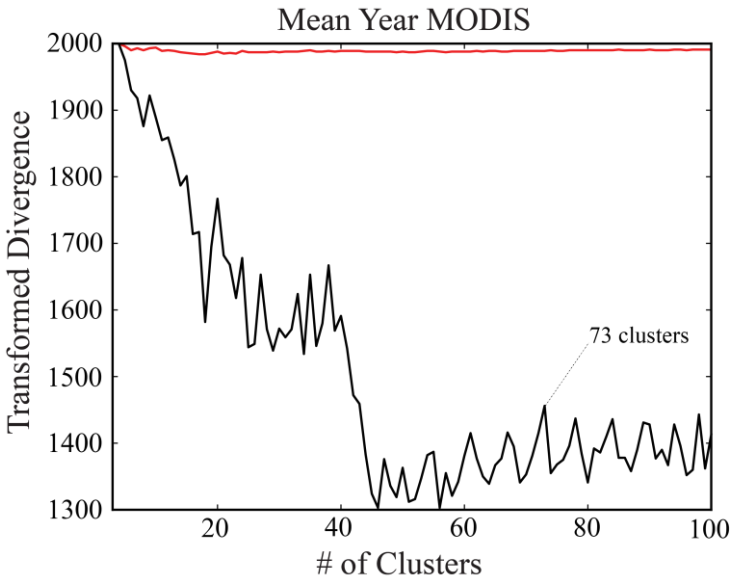


Figure 7. Transformed divergence (red) and minimum transformed divergence (black) as a function of number of clusters of mean year MODIS data reduction; these are used for for determining local maxima of the minimum transformed divergence. A 73-class scheme was ultimately selected as optimal.

Optimal ISODATA classification schemes for both the traditional and novel reduced-data time series were identified by iterating through k = 3, 4…100, where k is the desired number of output clusters as requested by ISODATA. The maximum number of classes was limited to 100 to ease interpretation while allowing for extraction of spatially small classes. Each proposed classification scheme resulting from a different ISODATA k was scrutinized using cluster validation techniques. For each execution, the transformed divergence index [35] was calculated using ERDAS Imagine. For both the traditional and novel branches of the overall workflow (Figure 5), an optimal classification scheme was determined manually using a combination of average transformed

divergence and minimum transformed divergence (Figure 7). This important *k* parameter adjustment and cluster validity procedure ensured that the dimensionally distinct traditional and novel clustering approaches were comparable, as each was based on an optimal number of output clusters requested.

3. Results and Discussion

3.1. Mean year Reduced Data

The structure of mean year reduced data is inflexible; it will always consist of a time series with the same number of rasters as annual compositing periods in the full time series, regardless of the number of years spanned. This characteristic trades flexibility of amount of data reduction for reduced computational complexity of the reduced data time series. Thus, the MODIS time series, which spanned 12 years, was reduced to 23 rasters (one for every 16 days in a year).

3.1.1. Transformed Divergence

Transformed divergence remained relatively high in all classification schemes for MODIS reduced mean time series (Figure 7). The minimum transformed divergence became useful in this scenario as selection criteria for optimal classification schemes. Local maxima of the calculated minimum transformed divergence indicated classification schemes that divide the data relatively well. The seven-class and 20-class schemes for the MODIS mean year reduced data time series did not divide the northern Fertile Crescent into small enough sub-areas. In the ultimately selected 73-class scheme, 11 of the classes were major contributors to the area of interest, and another 13 contributed meaningful spatial segregation. A total of 47 other classes intersected the study area but were very minor contributors or extremely scattered throughout the area. Only two classes did not intersect the study area at all. The 73-class scheme was also chosen for further development of a MODIS cluster map of the northern Fertile Crescent due to the reasonable expectation of interpreting the 24 major and minor contributing classes and the high level of spatial clustering apparent.

3.2. Segment Reduced Data

Selection strategy for the best segmentation scheme balanced the need to reduce the amount of data while maintaining appropriate spatial resolution (aggregation) or mean segment size (Figure 8). Percent data reduction is given more weight in the selection process because of the necessity to limit subsequent data processing costs. As the data reduction curve approaches 100%, the value returned by each successive segmentation scheme is reduced as shown by the flattening of the curve. Similarly, as the segmentation scale increases, the value returned by the average segment size is reduced due to the exponential growth of average segment size. An estimated starting point to eliminate segmentation schemes is at the transition of vertical trend to horizontal trend on the percent data reduction curve, specifically at scale 24. Consequently, schemes of scale 24 and below were eliminated from the candidate optimal segmentation scheme pool.

To further identify an optimal MODIS segmentation scheme, we limited the candidate segmentation scales to those which reduced the full time series to attain percent reduction comparable to that of the mean year time series (which reduced the full data time series by 92%). Consequently, the optimal number of segments in a segmentation scheme was at most 3,604,898, which eliminated scale 27 from consideration. Segment size increased dramatically with each successive segmentation scheme while the change in percent data reduction diminished rapidly. Scale 30 was selected as optimal due to it being the next smallest given the above constraints.

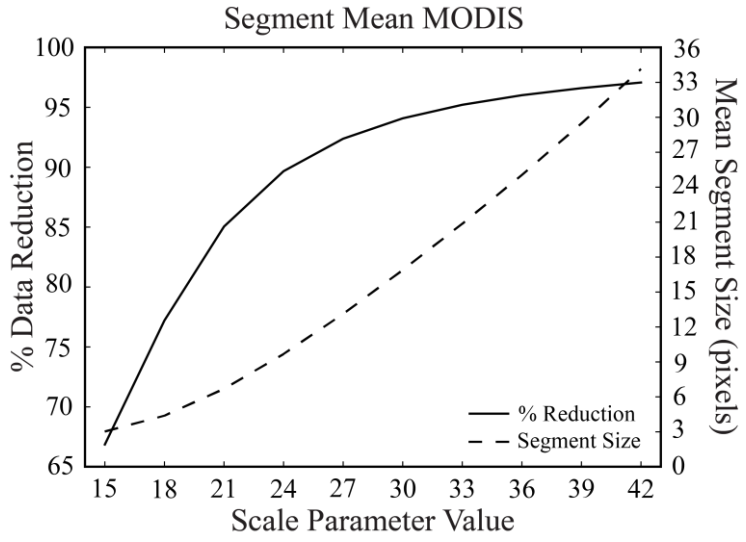


Figure 8. Data reduction and mean segment size for various segmentation scale parameters for segment mean MODIS time series. This graph was used to guide the selection of an optimal segmentation scheme.

3.2.1. Transformed Divergence

The ERDAS Imagine transformed divergence separability function did not produce reliable results for the segment mean time series. (The average and minimum transformed divergence for each classification scheme produced a value of 2000, which is the maximum value possible for this metric.) As an alternative, the same number of classes from the mean year classification schemes was used for the segment mean classification for the MODIS time series.

3.3. Comparison of Mean Year Classes to Segmented Classes

ISODATA classes (clusters) were categorized as major, minor, and scattered to aide interpretation of spatial distribution of each classification scheme (Table 2). Major classes contribute significant portions of the total study area, minor classes contributed smaller clusters of pixels, and scattered classes contributed only a few pixels per class. These three categories were used to guide the creation of classification maps for each data type and reduction method. As a result, only major and minor classes were included in the results (Figures 9–11).

Table 2. ISODATA classes (clusters) intersecting the study area. Major classes contributed significant portions of the total study area, and minor classes, comprised of spatially interesting zones, contributed smaller clusters of pixels. Scattered classes contributed only a few pixels per cluster in the study area.

Methodology	Major	Minor	Scattered
Traditional mean year	25, 26, 27, 29, 30, 41, 43, 46, 49, 56, 58	24, 28, 44, 47, 48, 55, 57, 59, 60, 61, 63, 64, 72	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 31, 32, 33, 34, 35, 37, 38, 39, 40, 42, 45, 50, 51, 52, 53, 54, 62, 65, 66, 67, 68, 69, 70, 71, 73
Novel segment mean	32, 34, 35, 36, 39, 41, 42	37, 38, 40	25, 26, 28, 29, 30, 33

Comparison of the MODIS classification schemes shows drastic differences (Figures 9–10). The deserts are more fractured into multiple classes in the mean year scheme. The upper study area was also simplified from many classes to only a few in the segment mean method. Most revealing of the change in classification schemes came from the class inclusion chart (Table 2). The number of major classes dropped from 11 to seven in the segment mean method, and minor classes dropped from 13 to only three. A driver of this drastic change was the reduced number of valid classes (with membership of at least one pixel) produced during the ISODATA clustering of the segment mean time series.

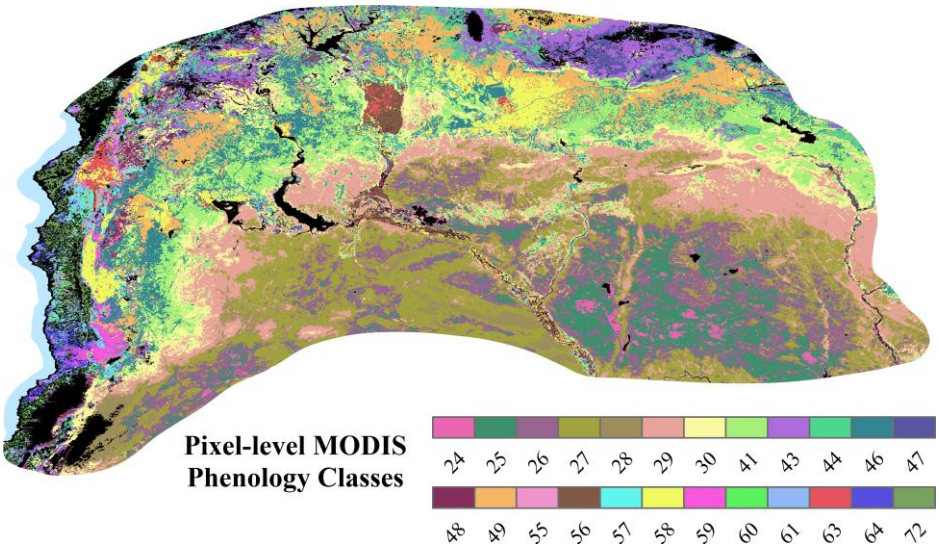


Figure 9. Major and minor classes from the MODIS mean year data reduction method.

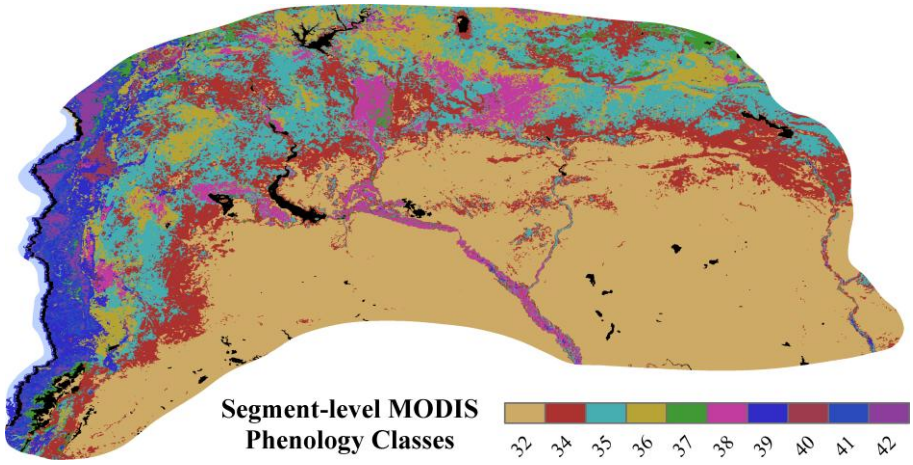


Figure 10. Major and minor classes from the MODIS segment mean data reduction method.

3.3.1. Spatial Clustering and Temporal Patterning of Classes

A hallmark difference between mean year and segment mean classes is the appearance of spatial smoothing similar to the effect of a low-pass filter. The segment mean method generated a reduced number of speckled classes (lone pixels or small groups of pixels of differing class membership within other classes), producing more homogeneous spatial clusters that better represent regional differences in land surface phenology. To aid interpretation of differences between results of mean year and segment mean data reduction methods for creation of individual classes, specific areas were selected for more detailed analysis. These areas were compared with other higher-resolution imagery datasets, including historical Landsat Thematic Mapper (TM) and Operational Land Imager (OLI) data and high spatial resolution aerial and satellite imagery freely available through Esri's ArcMap. This discussion focuses on areas that are characterized by intensive irrigation agriculture, something that is captured far better by the segment mean classification method than the traditional mean year approach.

Many parts of the study area have seen rather dramatic increases in irrigated agriculture since the 1970s, driven both by the creation of large reservoirs along river valleys as well as the introduction of mechanized pumps that tap deep groundwater [54]. Irrigation from both of these sources has enabled the introduction of new crops with high water demands, such as cotton, the extension of cultivation to areas not previously farmed, and the maturation of crops during the dry months of the late summer and early fall. These differences should be easily captured by time-series classification as they result in higher NDVI values generally, and appear at different times of year than under dry-farming conditions. Such areas of intensive irrigation are represented by mean year class 56 and segment mean class 38 (Figures 11–14).

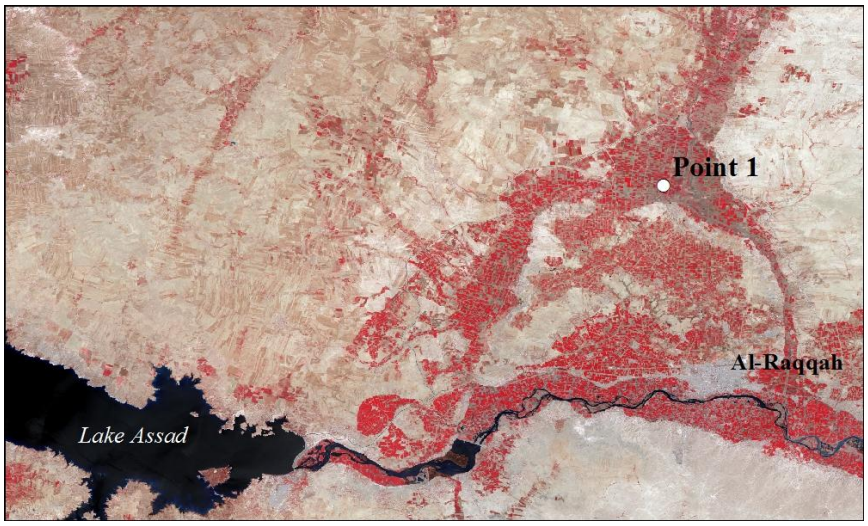


Figure 11. Point of interest (Point 1) near Al-Raqqah, Syria. The surrounding area is dominated by irrigated agriculture. Landsat 8 Operational Land Imager (OLI) imagery acquired 24 Apr 2013 (RGB = NIR, red, green) is in the background.

A good example of the differences represented by the two methods can be seen along the Middle Euphrates River in Syria, at the confluence of a tributary known as the Balikh River, near the modern war-torn city of Al-Raqqah. Traditionally, the region saw only limited irrigation immediately adjacent to the river valleys, and was otherwise dependent on dry-farmed cereals. However, the construction of the Tabqa Dam and the creation of Lake Assad enabled widespread intensification of irrigated agriculture since the 1970s [54]. Recent Landsat 8 OLI imagery clearly shows the extent of irrigated agriculture in the region (Figure 11). We selected a point of interest within the zone of irrigated agriculture (Point 1).

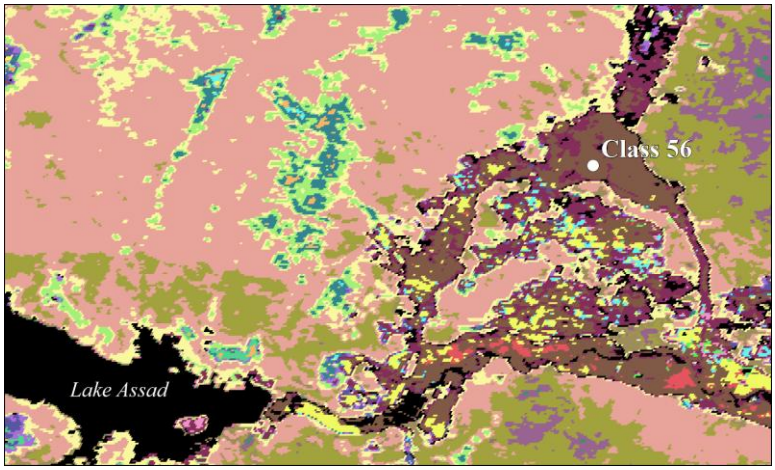


Figure 12. The point of interest intersects class 56 of the mean year classification scheme; this class appears to generally correspond to cropped areas (Figure 11).

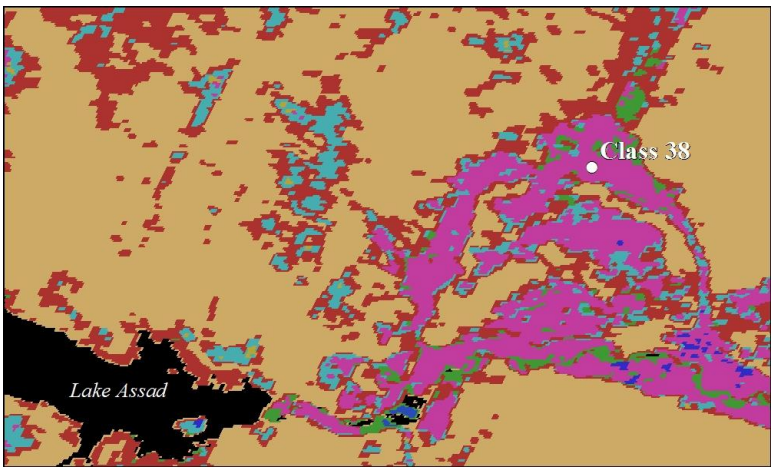


Figure 13. The point of interest intersects with the main agricultural class of the segment mean classification scheme, class 38; this class is less fragmented than the mean year class 56 (Figure 12).

Point 1 intersects class 56 of the mean year classification scheme, and while much of the surrounding irrigated areas are similarly classified, the region is dissected by many other minor classes appearing as small patches or isolated pixels, something that would complicate most analyses of resultant phenological classes. In contrast, the segment mean classification scheme, in which Point 1 is represented by class 38, shows a far more homogenous area that corresponds well to zones that appear on high-resolution imagery to be irrigated agriculture. Furthermore, the areas falling into class 38 are fringed by patches of zone 34, reflecting either dry-farmed or less-intensively irrigated agricultural areas. Thus, the segment mean classification method appears to function very well at quantitatively demarcating pixels with patterns of land use over time that are qualitatively similar. This suggests that the land surface phenology technique described may be valuable in land use classification workflows, especially where multi-temporal information content is critical in the extraction of land use (e.g., agricultural) classes

3.3.2. Temporal Trajectory of Selected Classes

The superiority of the segment mean classification method for discriminating meaningfully different patterns of agricultural land use is evident in the temporal trajectories of individual classes. Graphs of the temporal trajectories of individual classes, illustrating the mean, maximum, and minimum NDVI throughout the study period, were produced to provide insight into class structure.

Analysis of the temporal trajectory of mean NDVI values shows that both methods produced a cluster with double peaks, reflecting the presence of two annual cropping cycles. The double peak in these classes is driven by a practice of growing a first crop in the spring, with a harvest in the early summer, followed by a period of ploughing and a second planting in mid-summer, with a harvest in the fall. In the highly seasonal climate of the Middle East, with cool, wet winters and hot, dry summers, a cropping cycle with two annual harvests is generally only possible with the aid of irrigation. The mean year method reveals this double peak in NDVI. However, the period between crop cycles is significantly smoothed in the mean year cluster by comparison to the more well-defined cropping periods of the

segment mean cluster (Figure 14). Additionally, the peak NDVI cropping period is switched between the two classes. For the mean year class, the peak NDVI period comes after the first cropping period of each year, while the segment mean class shows a peak NDVI during the first cropping period of each year. Thus in addition to its better spatial representation of differing land use practices, the segment mean classification method also captures far more detail in the temporal characteristics of an individual cluster.

The segment mean method has an important advantage compared with methods used in other works that incorporated spatial object-based segmentation in conjunction with vegetation index time series [9,30–33]. Carefully parameterized image segmentation applied to the full time series ensures that regional or global products that contain homogenous landscape units provide a complete temporal signature for a growing number of land surface phenology-driven applications. Post-segmentation reduction of the time series can still address application-specific goals (e.g., fire ignition susceptibility versus ecosystem monitoring). Such reduction can still incorporate PCA or other approaches such as stepwise discriminant analysis, both tested in related pre-segmentation workflow designs [9,55]. A disadvantage to the segment mean approach is the added computational costs of segmenting a full time series, which can be significant for a global study area spanning multiple decades. (Segmentation in the Fertile Crescent study area required up to three hours on a GIS-class server.) However, the tiling and stitching approach demonstrated (Figure 6) using Trimble's eCognition Server does allow this approach to be scaled using existing geospatial infrastructure (including geoprocessing clusters). Furthermore, the proliferation of object-based techniques in high spatial resolution applications (e.g., that leverage data from unmanned aircraft systems or UAS), ensures that availability of such resources will likely increase.

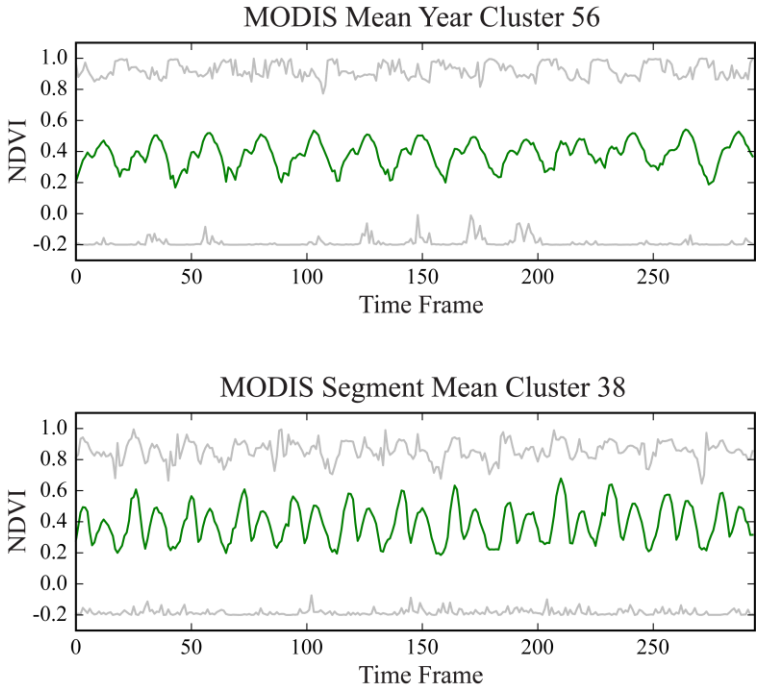


Figure 14. Cluster minimum, maximum, and mean (green) NDVI for clusters intersecting the point of interest throughout the original MODIS-derived NDVI time series.

With these results, additional work is still required to refine the segment mean data reduction method in at least three ways. First, work is needed to refine guidelines for quantitative determination of an optimal scale parameter for the spatial data reduction technique. Existing quantitative techniques [55] assume reference segments can be identified, which is problematic if each segment captures hundreds of image epochs. Second, additional research is required to determine an improved cluster validity index to use to select the optimal classification scheme. Although our study used a quantitative measure (the minimum transformed divergence) to guide classification scheme selection, work is needed to compare the effect of using different CVIs. Third, we applied segmentation on the full time series using a manual selection process based on investigative heuristics deemed useful for demonstration of the proposed technique. However, there may be a variety of pre-segmentation processing steps and/or adaptively tuned segmentation configuration parameters that could augment homogeneity in the resulting segments, and these should be carefully examined in a workflow analysis. In the context of comparable dimensionalities, advances in object-based hyperspectral image analysis [56,57] may prove helpful in this regard. Finally, future research is also needed to determine the best distance (similarity) measure for use in unsupervised classification in order to take advantage of the properties of the segment mean method. More effective similarity measures, which use the temporal information retained by the segment mean data reduction method, should take into account the shape of the vegetation index curve over time.

4. Conclusion

This regional to globally applicable study developed a method to successfully preserve temporal information while reducing data size in the classification of long, complex remote sensing-derived vegetation index time-series. In a northern Fertile Crescent case study, two land surface phenology classification maps were produced using a 250 m MODIS 12-year time series. Through spatial and time series analysis, important differences were identified between unsupervised classifications based on mean year reduced versus segment mean reduced long time series NDVI.

Each technique has advantages depending on the desired spatial or temporal granularity. The traditional mean year reduction method may be more suitable for analysis of intra-annual variation in land surface phenology, and is computationally more efficient. However, the segment mean method offers several key advantages. First, the segment mean approach, showing both spatial smoothing similar to a low-pass filter and a reduction of the number of total classes covering the study area, produces more homogeneous clusters that represent valuable agricultural land use information. Second, the method also preserves information regarding the temporal trajectories of classes that are not well represented in a traditional mean year data reduction approach. These results offer a valuable contribution to future work attempting to extract land surface phenology patterns from remote sensing-derived regional to global vegetation index time series.

Acknowledgements

This research was funded through NASA ROSES Space Archaeology award number NNX10AM39G, “Settlement Systems and Environmental Change in the Northern Fertile Crescent”.

Additional support was provided by AmericaView through USGS grant #G14AP00002. The authors directed the design, analysis, and preparation of published material and data.

Conflict of interest

All authors declare no conflict of interest in this paper.

References

1. Garonna I, de Jong R, Schaepman ME (2016) Variability and Evolution of Global Land Surface Phenology Over the Past Three Decades (1982–2012). *Glob Ch Biol* 22: 1456-1468.
2. Morisette JT, Richardson AD, Knapp AK, et al. (2009) Tracking the Rhythm of the Seasons in the Face of Global Change: Phenological Research in the 21st Century. *Front in Ecol and the Environ* 7: 253-260.
3. Trigo RM, Gouveia CM, Barriopedro D (2010) The Intense 2007–2009 Drought in the Fertile Crescent: Impacts and Associated Atmospheric Circulation. *Agric For Meteorol* 150: 1245-1257.
4. Bajgiran PR, Darvishsefat AA, Khalili A, et al. (2008) Using AVHRR-based Vegetation Indices for Drought Monitoring in the Northwest of Iran. *J of Arid Environ* 72: 1086-1096.
5. Beaumont P (1996) Agricultural and Environmental Changes in the Upper Euphrates Catchment of Turkey and Syria and Their Political and Economic Implications. *Appl Geogr* 16: 137-157.
6. Perrin de Brichambaut G, Wallen CC (1963) A Study of Agroclimatology in Semi-arid and Arid Zones of the Near East. Geneva, Switzerland: World Meteorological Organisation. Technical Note No. 56 Technical Note No. 56. p. 64.
7. Wilkinson TJ (2003) Archaeological Landscapes of the Near East. Tucson: University of Arizona Press.
8. Gu Y, Brown JF, Miura T, et al. (2010) Phenological Classification of the United States: A Geographic Framework for Extending Multi-Sensor Time-Series Data. *Remote Sens* 2: 526-544.
9. Bisquert M, Begue A, Deshayes M (2015) Object-based Delineation of Homogeneous Landscape Units at Regional Scale Based on MODIS Time Series. *Int J of Appl Earth Obs and Geoinform* 37: 72-82.
10. Gómez C, White JC, Wulder MA (2016) Optical Remotely Sensed Time Series Data for Land Cover Classification: A Review. *ISPRS J of Photogramm and Remote Sens* 116: 55-72.
11. Shao Y, Lunetta RS, Wheeler B, et al. (2016) An Evaluation of Time-series Smoothing Algorithms for Land-cover Classifications Using MODIS-NDVI Multi-temporal Data. *Remote Sens of Environ* 174: 258-265.
12. Bajocco S, Dragozi E, Gitas I, et al. (2015) Mapping Forest Fuels through Vegetation Phenology: the Role of Coarse-resolution Satellite Time-series. *Plos One* 10.
13. Jensen JR (2007) Remote Sensing of the Environment: An Earth Resource Perspective; Clarke KC, editor. Upper Saddle River, New Jersey: Prentice Hall. p. 592.
14. Holben BN (1986) Characteristics of Maximum-value Composite Images from Temporal AVHRR Data. *Int J of Remote Sens* 7: 1417-1434.

15. Vermote E, Kaufman YJ (1995) Absolute Calibration of AVHRR Visible and Near-infrared Channels Using Ocean and Cloud Views. *Int J of Remote Sens* 16: 2317-2340.
16. Los SO (1998) Estimation of the Ratio of Sensor Degradation Between NOAA AVHRR Channels 1 and 2 from Monthly NDVI Composites. *IEEE Trans on Geosci and Remote Sens* 36: 206-213.
17. Vermote E, El Saleous N, Kaufman YJ, et al. (1997) Data Pre-processing: Stratospheric Aerosol Perturbing Effect on the Remote Sensing of Vegetation: Correction Method for the Composite NDVI After the Pinatubo Eruption. *Remote Sens Rev* 15: 7-21.
18. Pinzón JE, Brown ME, Tucker CJ (2005) EMD Correction of Orbital Drift Artifacts in Satellite Data Stream. *Hilbert-Huang Transform: Introd and Appl*. pp. 167-186.
19. Clark RN, Swayze GA, Wise RA, et al. (2007) USGS Digital Spectral Library splib06a. Denver, Colorado: U.S. Geological Survey.
20. Tucker CJ, Pinzon JE, Brown ME, et al. (2005) An Extended AVHRR 8-km NDVI Data Set Compatible with MODIS and SPOT Vegetation NDVI Data. *Int J of Remote Sens* 26: 4485-4498.
21. Deering DW, Rouse Jr. JW, Haas RH, et al. (1975) Measuring 'Forage Production' of Grazing Units From Landsat MSS Data. *Proc of the 10th Int Symp on Remote Sens of the Environ* 2: 1169-1178.
22. Tucker CJ (1979) Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sens of Environ* 8: 127-150.
23. Geerken RA (2009) An Algorithm to Classify and Monitor Seasonal Variations in Vegetation Phenologies and their Inter-annual Change. *ISPRS J of Photogramm and Remote Sens* 64: 422-431.
24. Kouchoukos N (2001) Satellite Images and Near Eastern Landscapes. *Near Earstern Archaeol* 64: 80-91.
25. Nguyen TTH, De Bie CAJM, Ali A, et al. (2012) Mapping the Irrigated Rice Cropping Patterns of the Mekong Delta, Vietnam, Through Hyper-temporal SPOT NDVI Image Analysis. *Int J of Remote Sens* 33: 415-434.
26. Al-Bakri JT, Taylor JC (2003) Application of NOAA AVHRR for Monitoring Vegetation Conditions and Biomass in Jordan. *J of Arid Environ* 54: 579-593.
27. Benedetti R, Rossini P, Taddei R (1994) Vegetation Classification in the Middle Mediterranean Area by Satellite Data. *Int J of Remote Sens* 15: 583-596.
28. Evans J, Geerken R (2006) Classifying Rangeland Vegetation Type and Coverage Using a Fourier Component Based Similarity Measure. *Remote Sens of Environ* 105: 1-8.
29. Hölbling D, Friedl B, Eisank C (2015) An Object-based Approach for Semi-automated Landslide Change Detection and Attribution of Changes to Landslide Classes in Northern Taiwan. *Earth Sci Inform* 8: 327-335.
30. Bontemps S, Bogaert P, Titeux N, et al. (2008) An Object-based Change Detection Method Accounting for Temporal Dependences in Time Series with Medium to Coarse Spatial Resolution. *Remote Sens of Environ* 112: 3181-3191.
31. Hüttich C, Herold M, Strohbach B, et al. (2011) Integrating In-situ, Landsat, and MODIS Data for Mapping in Southern African Savannas: Experiences of LCCS-based Land-cover Mapping in the Kalahari in Namibia. *Environ Monit and Assess* 176: 531-547.

32. De Angelis A, Bajocco S, Ricotta C (2012) Phenological Variability Drives the Distribution of Wildfires in Sardinia. *Landsc Ecol* 27: 1535-1545.
33. Zhong LH, Gong P, Biging GS (2012) Phenology-based Crop Classification Algorithm and its Implications on Agricultural Water Use Assessments in California's Central Valley. *Photogramm Eng and Remote Sens* 78: 799-813.
34. Tou JT, Gonzalez RC (1974) Pattern Recognition Principles. Reading, Massachusetts: Addison-Wesley.
35. Swain PH, Davis SM (1978) Remote Sensing: The Quantitative Approach: McGraw-Hill International Book Co.
36. Cohen J (1960) A Coefficient of Agreement for Nominal Scales. *Educ Psychol Meas* 20: 37-46.
37. Jensen JR (2016) Introductory Digital Image Processing: A Remote Sensing Perspecitve; Clarke KC, editor: Pearson Education. p. 623.
38. Arbelaitz O, Gurrutxaga I, Muguerza J, et al. (2013) An Extensive Comparative Study of Cluster Validity Indices. *Pattern Recognit* 46: 243-256.
39. ERDAS (2010) ERDAS Field Guide. Norcross, GA: ERDAS, Inc. 812 p.
40. Ward JH (1963) Hierarchical Grouping to Optimize an Objective Function. *J Am Stat Assoc* 58: 236-244.
41. Baatz M, Schäpe A. Multiresolution Segmentation: an Optimization Approach for High Quality Multi-scale Image Segmentation; 2000 2000. Herbert Wichmann Verlag.
42. Sun W, Li W, Li J, et al. (2015) Band Selection Using Sparse Nonnegative Matrix Factorization with the Thresholded Earth's Mover Distance for Hyperspectral Imagery Classification. *Earth Sci Inform* 8: 907-918.
43. Hole F, Zaitchik BF (2007) Policies, Plans, Practice, and Prospects: Irrigation in Northeastern Syria. *Land Degrad & Dev* 18: 133-152.
44. Riehl S (2012) Variability in Ancient Near Eastern Environmental and Agricultural Development. *J of Arid Environ* 86: 113-121.
45. Wilkinson TJ (1994) The Structure and Dynamics of Dry-Farming States in Upper Mesopotamia. *Curr Anthropol* 35: 483-520.
46. Staubwasser M, Weiss H (2006) Holocene Climate and Cultural Evolution in Late Prehistoric–early Historic West Asia. *Quat Res* 66: 372-387.
47. Bar-Yosef O (2011) Climatic Fluctuations and Early Farming in West and East Asia. *Curr Anthropol* 52: S175-S193.
48. Ur JA (2010) Cycles of Civilization in Northern Mesopotamia, 4400-2000 BC. *J of Archaeol Res* 18: 387-431.
49. NASA (2016) Reverb | ECHO: The Next Generation Earth Science Discovery Tool.
50. Solano R, Didan K, Jacobson A, et al. (2010) MODIS Vegetation Index User's Guide (MOD13 Series) Version 2.00. Vegetation Index and Phenology Lab, The University of Arizona.
51. USGS (2016) USGS Global Visualization Viewer.
52. Trimble (2013) eCognition Software.
53. Congalton R (2010) Remote Sensing: An Overview. *GISci & Remote Sens* 47: 443-459.
54. Casana J, Cothren J, Kalayci T (2012) Swords into Ploughshares: Archaeological Applications of CORONA Satellite Imagery in the Near East. *Internet Archaeology*.

55. Li QT, Wang CZ, Zhang B, et al. (2015) Object-based Crop Classification with Landsat-MODIS Enhanced Time-series Data. *Remote Sens* 7: 16091-16107.
56. Zehtabian A, Ghassemian H (2015) An Adaptive Pixon Extraction Technique for Multispectral/Hyperspectral Image Classification. *IEEE Geosci and Remote Sens Lett* 12: 831-835.
57. Zehtabian A, Ghassemian H (2016) Automatic Object-Based Hyperspectral Image Classification Using Complex Diffusions and a New Distance Metric. *IEEE Trans on Geosci and Remote Sens* 54: 4106-4114.

AIMS Press