Remote Sensing of Environment 158 (2015) 478–491

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Improved time series land cover classification by missing-observation-adaptive nonlinear dimensionality reduction

L. Yan *, D.P. Roy

Geospatial Science Center of Excellence, South Dakota State University Brookings, SD 57007, USA

ARTICLE INFO

Article history:
Received 11 June 2014
Received in revised form 12 November 2014
Accepted 15 November 2014
Available online 12 December 2014

Keywords:
Landsat
Land cover classification
Time series
Nonlinear dimensionality reduction
WELD

ABSTRACT

Dimensionality reduction (DR) is a widely used technique to address the curse of dimensionality when high-dimensional remotely sensed data, such as multi-temporal or hyperspectral imagery, are analyzed. Nonlinear DR algorithms, also referred to as manifold learning algorithms, have been successfully applied to hyperspectral data and provide improved performance compared with linear DR algorithms. However, DR algorithms cannot handle missing data that are common in multi-temporal imagery. In this paper, the Laplacian Eigenmaps (LE) nonlinear DR algorithm was refined for application to multi-temporal satellite data with large proportions of missing data. Refined LE algorithms were applied to 52-week Landsat time series for three study areas in Texas, Kansas and South Dakota that have different amounts of missing data and land cover complexity. A series of random forest classifications were conducted on the refined LE DR bands using varying proportions of training data provided by the United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) Cropland Data Layer (CDL); these classification results were compared with conventional metrics-based random forest classifications. Experimental results show that compared with the metrics approach, higher per-class and overall classification accuracies were obtained using the refined LE DR bands of multispectral reflectance time series, and the number of training samples required to achieve a given degree of classification accuracy was also reduced. The approach of applying the refined LE to multispectral reflectance time series is promising in that it is automated and provides dimensionality-reduced data with desirable classification properties. The implications of this research and possibilities for future algorithm development and application are discussed.

1. Introduction

There is an established heritage for the application of dimensionality reduction (DR) techniques to multispectral satellite time series prior to land cover classification. DR techniques transform high-dimensional data into data with lower dimensions that ideally maximize the information content and minimize noise (Bellman, 2003; Hughes, 1968). For example, the principal component analysis (PCA) DR technique has been applied to satellite time series to provide new dimensionality-reduced bands for land cover and land cover change classification (Collins & Woodcock, 1996; Murthy, Raju, & Badrinath, 2003; Small, 2012; Townshend, Justice, & Kalb, 1987; Zhong & Wang, 2006). PCA performs a linear mapping of the data to a lower dimensional space so that the variance of the transformed data is maximized. A number of other DR techniques based on linear transformations have been developed including linear discriminant analysis (LDA) (Fisher, 1936; Martinez & Kak, 2001), projection pursuit (PP) (Friedman & Tukey, 1974; Jimenez & Landgrebe, 1999), minimum noise fraction (MNF) (Green, Berman, Switzer, & Craig, 1988), independent component analysis (ICA) (Hyvärinen, 1999; Hyvärinen & Oja, 2000), and spatial–spectral eigenvector derivation (SSEVD) (Rogge, Bachmanna, Rivard, Nielsen, & Feng, 2014). However, linear DR techniques do not accommodate the intrinsic nonlinear characteristics of optical wavelength remotely sensed data, whereby the remotely sensed contribution of the observed scene components are not linearly proportional to their surface areas, and that are particularly apparent for vegetation. Nonlinearity is introduced by multiple scattering between different scene components that varies as a function of the wavelength, the viewing and illumination geometry, and the three dimensional structure of the scene components (Schaaf et al., 2002; Somers et al., 2009). Spectral indices based on ratios of reflectance, such as the normalized difference vegetation index (NDVI), are not linear functions of reflectance and so also have nonlinear characteristics (Verstraete & Pinty, 1996), and, for example, their use in spectral linear unmixing is not recommended (Busetto, Meroni, & Colombo, 2008; Settle & Campbell, 1998). The reflectance contribution of most scene components change temporally and this may also introduce nonlinearity when time series data are considered. Nonlinear DR techniques have been used by the remote sensing community

* Corresponding author.
E-mail addresses: lin.yan@sdstate.edu (L. Yan), david.roy@sdstate.edu (D.P. Roy).

http://dx.doi.org/10.1016/j.rse.2014.11.024

predominantly for application to single date hyperspectral data, and have been found to provide improved performance for classification, target discrimination and end-member extraction, when compared with linear DR techniques (Bachmann, Ainsworth, & Fusina, 2005; Feilhauer, Faude, & Schmidtlein, 2011; Han & Goodenough, 2005; Zhang, Zhang, Tao, & Huang, 2013). A few studies have applied nonlinear DR techniques to single date multispectral satellite data (Journaux, Foucherot, & Gouton, 2006). However, nonlinear DR techniques have not been applied to multispectral satellite time series.

An established limitation of linear and nonlinear DR techniques is that they cannot handle missing data (Hubert, Rousseeuw, & Verboven, 2002). This is particularly problematic for satellite time series application, as missing data occurs frequently due to cloud obscuration (Ju & Roy, 2008) and also sometimes due to issues associated with the satellite acquisition, satellite to ground station data transmission, and data production errors (Roy, Lewis, Schaaf, Devadiga, & Boschetti, 2006; Roy et al., 2002). The impacts of missing data may be minimized in coarse spatial resolution time series by processing them into reduced temporal resolution time series. This has been achieved either by application of per-pixel temporal compositing procedures that select a best pixel observation every *n*-days (Holben, 1986) or by inversion of *n*-days of observations against a model of the surface bidirectional reflectance distribution function (BRDF) to estimate reflectance at consistent viewing and illumination angles (Schaaf et al., 2002). These techniques are less appropriate, however, for application to time series with lower temporal resolution, such as Landsat. Per-pixel temporal compositing procedures are difficult to implement reliably with Landsat data because of the low cloud-free observation frequency relative to surface changes (Roy et al., 2010; White et al., 2014) and Landsat BRDF inversion approaches do not work reliably because of the narrow Landsat field of view which precludes sampling of the intrinsic reflectance anisotropy of most land surfaces (Roy, Wulder, et al., 2014; Shuai, Masek, Gao, Schaaf, & He, 2014).

The current state of the practice for large area multi-temporal land cover classification is to derive metrics from the time series and then classify the metrics bands with a supervised (i.e., training data dependent) non-parametric classification approach. The choice of metrics, usually the maximum and quartile values of spectral indices and spectral bands over the time series, has been justified empirically in terms of attempting to capture seasonal land cover class spectral variations in a way that is robust to missing data (Broich et al., 2011; DeFries, Hansen, & Townshend, 1995; Friedl et al., 2010; Hansen et al., 2011, 2014). The nonlinearity in the metrics is assumed implicitly to be handled by the statistical classifier, which may be reasonable when non-parametric classifiers are used. However, while providing a practical and useful form of data reduction, the metrics-based classification approaches may not be optimal as compared with linear and nonlinear DR techniques, because they may use less of the temporal information available in the satellite time series.

In this paper, a recent nonlinear DR method developed for application to hyperspectral data is refined for application to multispectral satellite time series that have missing observations. The methodology can be applied to any satellite time series but is demonstrated in this study using Landsat data. Landsat data have an established and rapidly evolving heritage for multi-temporal classification of land cover, change, and disturbance (Hansen & Loveland, 2012; Roy, Wulder, et al., 2014), but often have missing data due to cloud obscuration (Kovalskyy & Roy, 2013) and, in historical data, due to satellite reception and acquisition issues (Loveland & Dwyer, 2012). The Landsat data and study areas are first described, followed by a description and justification for selection of the Laplacian Eigenmaps (LE) nonlinear DR method (Belkin & Niyogi, 2002) and how it was refined to handle missing data. Experiments are conducted in three study areas in Texas, Kansas, and South Dakota that encompass complex agricultural landscapes with differing amounts of missing Landsat data. Supervised random forest classification of the output of the refined DR method applied to 52 weeks of Landsat time series are compared with supervised random forest classification of the conventional metrics derived from the Landsat data. The classification accuracies of the LE and the metrics-based approaches generated using varying proportions of training data, defined by sampling United States Department of Agriculture (USDA) Cropland Data Layer (CDL) data, are examined, and implications and recommendations for future research are discussed.

2. Data

2.1. Landsat data

Web Enabled Landsat Data (WELD) Version 1.5 products were obtained from the USGS National Center for Earth Resources Observation and Science (EROS) (http://e4ftl01.cr.usgs.gov/WELD/). The products are defined in the Albers Equal Area conic projection in separate geolocated tiles of 5000 × 5000 30 m pixels and because they are temporally aligned, they are straightforward to use for time series classification applications. For example, they were used to generate 30 m conterminous United States (CONUS) annual land cover (Hansen et al., 2011) and 5-year land cover change (Hansen et al., 2014) classifications. Weekly, monthly, seasonal and annual WELD products are generated by application of a temporal compositing scheme to select a single best pixel observation for each reporting period (Roy et al., 2010). The annual product is generated by compositing 52 weeks of Landsat data, and the seasonal, monthly and weekly products are generated by compositing the Landsat data acquired in each season, month, or week, respectively. The version 1.5 WELD products were generated from every available Landsat 7 ETM+ Level 1 T processed image in the USGS Landsat archive with cloud cover ≤80%. Standard Level 1 T processing includes radiometric correction, systematic geometric correction, precision correction using ground control chips, and the use of a digital elevation model to correct parallax error due to local topographic relief, with a CONUS geolocation error less than 30 m (Lee, Storey, Choate, & Hayes, 2004).

In this study WELD products for climate year 2010 were used. The annual (December 2009 to November 2010), Summer (June to August 2010), Autumn (September to November 2010), seven monthly (April through October 2010), and 52 weekly products (1st December 2009 to November 30th 2010) were used. Their specific usage is described in Section 5.

All the WELD products store for each 30 m pixel location the six reflective top of atmosphere Landsat 7 Enhanced Thematic Mapper Plus (ETM+) bands, the two top of atmosphere thermal bands, bit packed band saturation information, Normalized Difference Vegetation Index (derived as the near-infrared minus the red reflectance divided by their sum), two cloud masks, the day of the year that the pixel value was sensed on, and the number of Landsat observations considered in the product period (week, month, season, or annual) (Roy et al., 2010). In this study, the Landsat ETM+ reflective wavelength bands 2 (green, 0.53–0.61 μm), 3 (red: 0.63–0.69 μm), 4 (near-infrared: 0.78–0.90 μm), 5 (middle-infrared: 1.55–1.75 μm), 7 (middle-infrared: 2.09–2.35 μm) and the NDVI were used. The shortest wavelength Landsat ETM+ band 1 (blue: 0.45–0.52 μm) was not used because it is overly sensitive to atmospheric scattering (Ju, Roy, Vermote, Masek, & Kovalskyy, 2012; Roy, Qin, et al., 2014). All WELD pixel values flagged as cloudy in both the cloud masks were removed as they were highly likely to be cloud contaminated (Roy et al., 2010).

2.2. Cropland Data Layer

The United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS) Cropland Data Layer (CDL) for 2010 was obtained from the CDL web site (http://nassgeodata.gmu.edu/CropScape/). The CDL data were used as a source of supervised classification training and test data. The CDL is generated annually using moderate resolution satellite imagery and extensive agricultural ground

truth via a supervised non-parametric classification approach (Boryan, Yang, Mueller, & Craig, 2011; Johnson & Mueller, 2010). The CDL defines about 110 land cover and crop type classes at 30 m (Johnson & Mueller, 2010). For 2010 the overall CONUS CDL classification accuracy is reported as 84.3% and the major field crops have 85% to 95% classification accuracies (Johnson, 2013). The CDL is defined in the same Albers Equal Area conic projection as the Version 1.5 WELD data.

3. Study areas

Study areas in Texas, Kansas, and South Dakota that encompass a variety of crop types, crop seasonality, fallow and idle cropland, grasses and pasture land, and degrees of missing Landsat data were selected. These study areas are complex to classify not only because they include natural vegetation covers but also because of within-field spectral variability associated with variations in soil moisture, salinity, fertility and nutrient limitations, pesticide, herbicide and fertilizer treatment, pollution, pests and diseases, and because of the temporal variability and spectral similarity between crops and non-crops as a function of their phenological stage, degree of soil background, and the time of satellite observation (Chang, Hansen, Pittman, Carroll, & DiMiceli, 2007; Hall & Badhwar, 1987; Johnson, 2013, 2014; Rao, 2008; Yan & Roy, 2014; Zhang et al., 2003).

Study areas composed of approximately 100,000 to 125,000 30 m pixels were extracted from the WELD and CDL data. Larger areas were not used because the computation required for the Laplacian Eigenmaps DR method increases rapidly with the spatial image dimensions (Section 4). In order to capture a greater diversity of crop types, the 30 m WELD and CDL data in the study areas were subsampled by selecting every 10th 30 m pixel east–west and north–south, respectively. Figs. 1 to 3 illustrate the 2010 CDL data of the three study areas and are shown using the standard color legends provided by the USDA NASS (http://www.nass.usda.gov/research/Cropland/SARS1a.htm). Table 1 summarizes the three study areas and the primary 2010 CDL classes within each study area.

Table 2 summarizes the percentage of valid (non-missing, not cloudy) 2010 weekly WELD 30 m observations for the three study areas. The number of weeks (n) with at least one valid 30 m pixel in each study area is summarized, which shows there can be many weeks with no valid observations at all, particularly for the South Dakota study area. The percentage of valid weekly WELD 30 m pixels is computed in two ways: as the sum of the number of valid 30 m pixels in each weekly subset divided by the product of n and the study area spatial dimensions; and also as the sum of the number of valid 30 m pixels in each weekly subset divided by the product of 52 and the study area spatial dimensions. The variation in n among the study areas is predominantly due to cloud cover which can have significant temporal and geographic variability across the conterminous United States at the time of Landsat 7 ETM+ overpass (Ju & Roy, 2008). Missing data also occur because, at mid-latitudes, a location is never overpassed by the Landsat 7 ETM+ more than once per week and sometimes it is not overpassed at all due to the satellite orbit and sensing geometry. Further, about 22% of the acquired ETM+ data are missing due to the Landsat 7 ETM+ scan line corrector (SLC) that failed in 2003 and reduced the usable data in each L1T scene by about 22% (Markham, Storey, Williams, & Irons, 2004). Over the 52 weeks, only about 30% of the study area data were valid (i.e. not missing or cloudy), but when considering only those weeks with at least one valid 30 m pixel over the study area, the percentage is increased by about 4% (Texas), 8% (Kansas) and 22% (South Dakota).

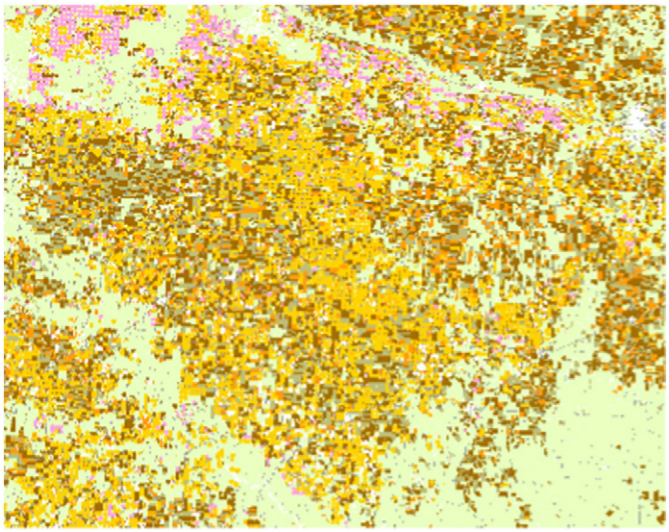

Fig. 2. Kansas 2010 CDL. 400 × 316 pixels subsampled from a 120 × 94.8 km region with predominant CDL classes of grass/pasture (35.2%), corn (18.6%), winter wheat (17.8%), fallow/idle cropland (11.6%), sorghum (6.7%), developed/open space (3.4%), alfalfa (3.1%). White shows the pixel locations where the CDL classes correspond to ≤2% of the study area.

The temporal variability of the WELD data for the CDL study area classes varies considerably. Fig. 4 shows the mean weekly Normalized Difference Vegetation Index (NDVI) temporal profiles for each CDL class in the study areas. Variations among the three study areas for the same crop class occur due to the differences in geographic locations, weather and site conditions, and planting and harvesting dates. The NDVI profiles have unimodal distributions with pronounced spring NDVI peaks for the winter wheat class and much longer growing

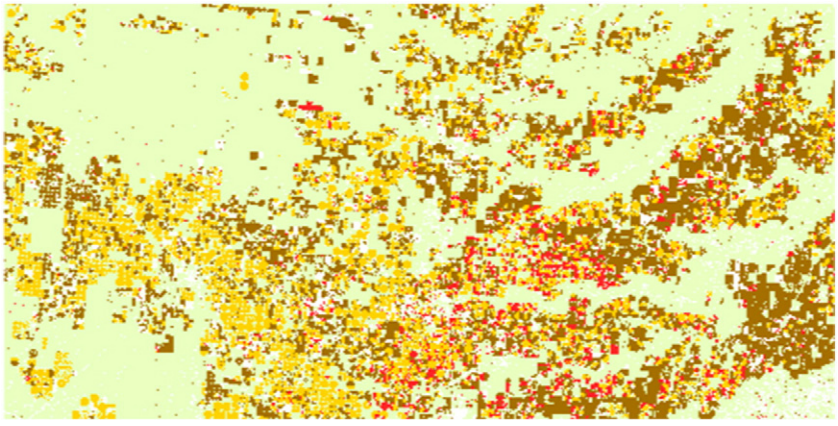

Fig. 1. Texas 2010 Cropland Data Layer (CDL), 500 × 250 30 m pixels subsampled from a 150 × 75 km region with predominant CDL classes of grass/pasture (55.7%), winter wheat (20.2%), corn (13.1%), cotton (3.3%). White shows the pixel locations where the CDL classes correspond to ≤2% of the study area.

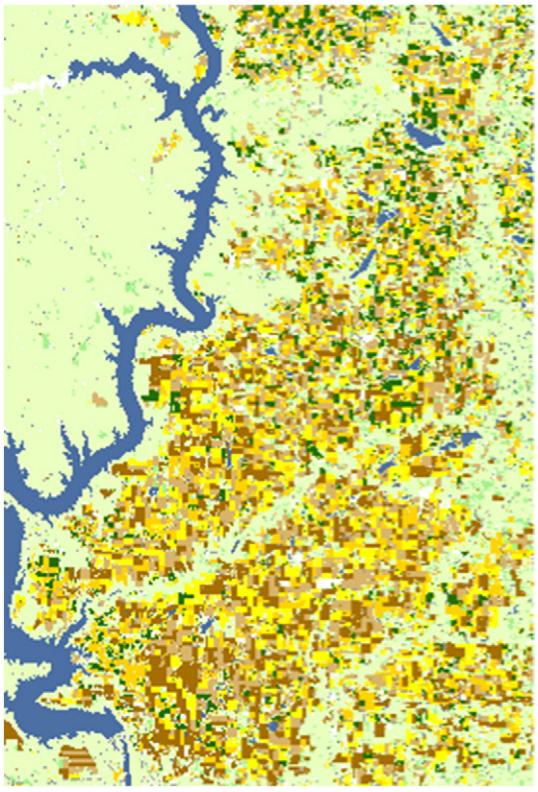

Fig. 3. South Dakota 2010 CDL. 270 × 400 pixels subsampled from an 81 × 120 km region with predominant CDL classes of grass/pasture (44.2%), winter wheat (10.1%), spring wheat (9.9%), corn (8.5%), sunflower (7.3%), open water (7.1%), soybeans (4.6%), hay (3.5%), developed/open space (2.1%). White shows the pixel locations where the CDL classes correspond to ≤2% of the study area.

seasons for the CDL classes of fallow/idle cropland, grass/pasture and hay. In the South Dakota study area, the sunflower and soybeans classes have similar NDVI profiles and this illustrates a fundamental problem for crop discrimination when only NDVI time series are used.

4. Laplacian Eigenmaps (LE) dimensionality reduction

4.1. Conventional LE dimensionality reduction

The Laplacian Eigenmaps (LE) DR method was chosen for this research for the following reasons. First, it is a nonlinear DR algorithm that has been demonstrated to provide better performance compared with linear DR algorithms with respect to application to hyperspectral data (Bachmann et al., 2005; Han & Goodenough, 2005; Zhang et al., 2013). Second, it is locality-preserving, which enables the information pertaining to different classes to be enhanced in separate DR bands (Yan & Niu, 2014). This characteristic is particularly attractive for land cover classification as it provides an approach to transform high-dimensional data into data with lower dimensions in a way that maximizes among-class separability. This characteristic could also be potentially used for land cover mapping based on a single DR band as has been suggested using the independent component analysis (ICA) linear DR method (Ozdogan, 2010). Third, because LE is a spectral-graph-based approach, it is simple to incorporate different distance metrics compared with other locality-preserving methods (Yan & Niu, 2014), which provides a means, with refinement, to handle missing data.

The LE DR algorithm uses a manifold learning approach to find a low-dimensional representation of high-dimensional data while preserving local manifold properties (Belkin & Niyogi, 2002). It originated from spectral graph theory, in which the properties of a neighborhood graph can be represented by the eigenvalues and eigenvectors of the graph's Laplacian matrix (Chung, 1996). Conventionally, satellite data are considered to be composed of n bands, and a single pixel can be described as a point in n-dimensional space, often termed feature space. The manifold structures in feature space can be represented by a neighborhood graph where every pixel is a graph node and is connected to its k nearest neighbors in feature space based on a distance metric, usually the Euclidean distance, and the connections between nodes have an attribute that describes their separation defined using the distance metric (Bachmann, Ainsworth, & Fusina, 2006; Bachmann et al., 2005; van der Maaten, Postma, & van den Herik, 2009). Different manifold learning methods have been developed to obtain a lower-dimensional representation of the original data from the neighborhood graph. The LE method achieves this by deriving the Laplacian eigenvectors, which are the dimensionality-reduced data, from the Laplacian matrix of the graph (Belkin & Niyogi, 2002) using the following process:

a) Read the satellite data composed of $ncol \times nrow$ pixels and n bands. For single date imagery (e.g., hyperspectral data), n is the number of spectral bands, and for multi-temporal data n corresponds to the product of the number of spectral bands and the number of images in the time series.
b) Calculate the m^2 pairwise distances among the m pixels ($m = ncol \times nrow$) in the n-dimensional feature space using the Euclidean distance defined as:

$$Euc(a,b) = \sqrt{\sum_{z=1}^{n}\left(x_z^a - x_z^b\right)^2} \quad (1)$$

where a and b are two pixels whose feature space values are defined by $x_{z\in[1\ldots n]}$. Then find the k nearest neighbors of each pixel in feature space by searching through the Euclidean distances. Next, construct a neighborhood graph stored as a sparse matrix G with dimensions of $m \times m$, and the element values in G are calculated by a Gaussian kernel function defined as:

$$g_{i,j} = e^{-\frac{Euc(a,b)}{2\sigma^2}} \quad (2)$$

Table 1
Summary of the three study area geographic characteristics and primary Crop Data Layer (CDL) classes.

Study area name	Longitudinal extent	Latitudinal extent	Area (km^2)	Number of subsampled 30 m pixels	Primary CDL 2010 classes (>2% study area proportion)
Texas (northern High Plains)	101.2070275°W to 102.9403276°W	35.8955492°N to 36.4780778 N°	11,250	500 × 250	Cotton (3.3%), corn (13.1%), winter wheat (20.2%), grass/pasture (55.7%)
Kansas (south-west Kansas)	99.9269083°W to 101.3481360°W	37.1080080°N to 37.9004428°N	11,376	400 × 316	Alfalfa (3.1%), developed/open space (3.4%), sorghum (6.7%), fallow/idle cropland (11.6%), winter wheat (17.8%), corn (18.6%), grass/pasture (35.2%)
South Dakota (central, along Missouri River)	99.5959749°W to 100.6848613°W	44.4043746°N to 45.4473078°N	9720	270 × 400	Developed/open space (2.1%), hay (3.5%), sorghum (4.6%), open water (7.1%), sunflower (7.3%), corn (8.5%), spring wheat (9.9%), winter wheat (10.1%), grass/pasture (44.2%)

Table 2
Percentage of valid weekly WELD pixel observations (52 weeks from 1st December 2009 to November 30th 2010).

Study Area	Percentage of valid weekly WELD 30 m pixel observations over the 52 weeks	Number of weeks with at least one valid 30 m pixel (n) in the study area	Percentage of valid weekly WELD 30 m pixel observations computed over the n weeks
Texas	30.3%	46	34.3%
Kansas	31.8%	41	38.4%
South Dakota	31.7%	31	53.2%

where σ controls the flatness of the kernel, i and j are indices (ranging from 1 to m) to pixels a and b, respectively, and $g_{i,j}$ is an element (i,j) of the matrix G that indicates a is one of the k nearest neighbors of b in feature space according to the Euclidean distance, and thus b is connected to a in the neighborhood graph.

c) Compute a diagonal matrix D whose entries are the row sums of G with $d_{i,i} = \sum_{j=1}^{m} g_{i,j}$ and compute the Laplacian matrix L with dimensions of $m \times m$ as:

$$L = D - G. \quad (3)$$

d) Solve the eigenvector:

$$Lv = \lambda Dv \quad (4)$$

where v are the eigenvectors and λ are the corresponding eigenvalues. The o smallest non-zero eigenvalues are used to select the o dimensionality-reduced data bands defined by the corresponding eigenvectors. For this study o was set to 20 (for reasons that are described in the analysis methodology section). Each dimensionality-reduced data band, i.e. the eigenvector, is composed of $ncol \times nrow$ real values.

The above process requires a large number of floating point operations. The main computational load is in the solution of Eq. (4) that has a complexity of $O((ncol \times nrow)^3)$ (Yan, Huang, & Jordan, 2009), i.e. the computational requirements increase rapidly with the satellite data spatial dimensions but not with the number of bands — this is a useful property when a large number of bands such as those found with multi-temporal or hyperspectral data are considered. However, for application over large spatial areas, this is a serious constraint that is discussed further in the paper conclusion.

4.2. Refined LE dimensionality reduction

Two versions of the LE nonlinear DR methodology were developed, both using the spectral angle mapper (SAM) to reduce sensitivity to missing data. For convenience, we refer to these two methods as LE-SAM and LE-SAM-R.

4.2.1. Spectral angle mapper (SAM)

The SAM is conventionally used to determine the spectral similarity between two pixels by calculating the angle subtended between their points in feature space and the feature space origin (Keshava, 2004; Kruse et al., 1993), which is defined as:

$$SAM(a,b) = \frac{\sum_{z=1}^{n} x_z^a x_z^b}{\sqrt{\sum_{z=1}^{n} \left(x_z^a\right)^2} \sqrt{\sum_{z=1}^{n} \left(x_z^b\right)^2}} \quad (5)$$

where a and b are two pixels whose feature space values are defined by $x_{z\in[1\ldots n]}$, and n (≥ 2) is the number of feature space dimensions, i.e. the number of spectral bands in a single date image, or the product of the number of spectral bands and the number of images in time series. The SAM is bounded in the range [0, 1]; when two pixels are identical, the SAM is equal to one. If specific feature space dimensions values for either pixels a or b are missing, the SAM can still be computed by ignoring the missing dimensions provided that the remaining number of dimensions is greater than one. The SAM has several attractive properties compared with the Euclidean distance. It is invariant to multiplicative scaling, i.e. when applied to calibrated reflectance, it is insensitive to exogenous reflectance brightness variations, for example, due to solar

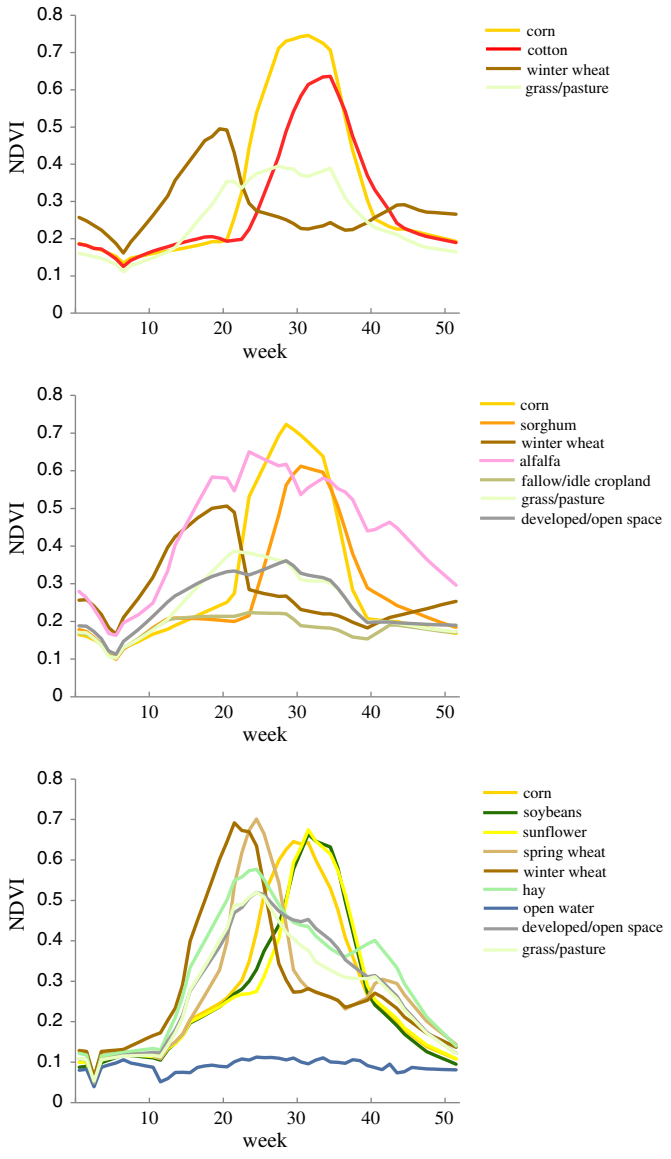

Fig. 4. Mean weekly NDVI temporal profiles of the Cropland Data Layer (CDL) classes for the three study areas, (top) Texas, (middle) Kansas, (bottom) South Dakota. Only profiles for CDL classes covering > 2% of the pixels in each study area (Figs. 1–3) are shown. Week 1 is the first week of December 2009 and week 52 is the last week of November 2010.

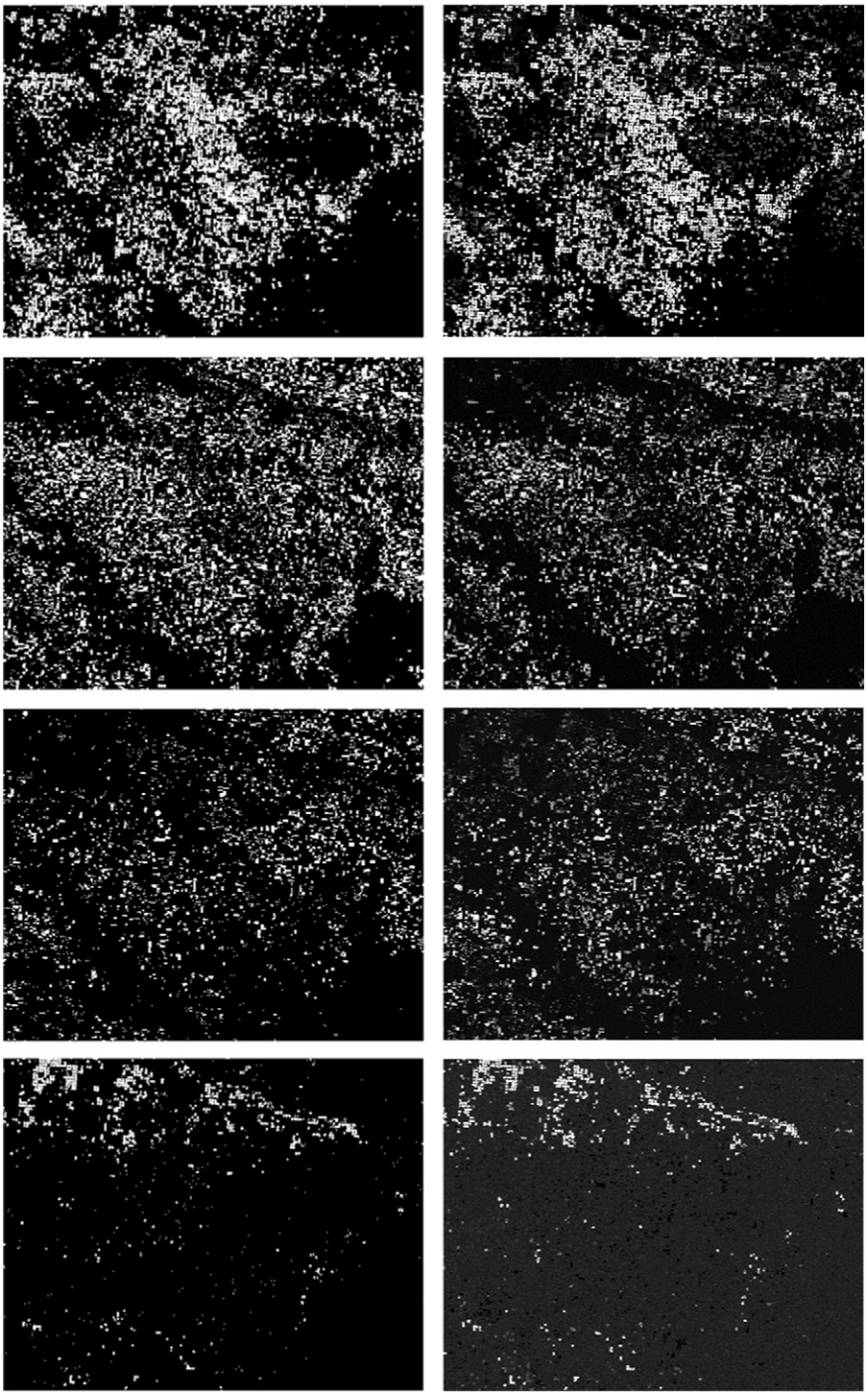

Fig. 5. Left: Kansas 30 m CDL classifications (Fig. 2) showing (from top to bottom) the four major agricultural crops of corn, winter wheat, sorghum and alfalfa in binary images. Right: the four selected LE-SAM-R DR bands (obtained from the 5-band reflectance weekly time series) that have the highest qualitative visual correspondence with the corresponding CDL classes in row.

geometry effects (Keshava, 2004). The idea of using SAM rather than a Euclidean metric for nonlinear dimensionality reduction was first described by Bachmann et al. (2005) who used it in the isometric mapping (ISOMAP) approach (Tenenbaum, de Silva, & Langford, 2000) applied to hyperspectral data; they noted the utility of SAM for mitigation of the effects of variable illumination, although SAM can be sensitive to low signal to noise over naturally dark and certain shaded targets. In addition, the SAM also has a non-monotonicity property whereby the SAM value will not necessarily increase with n; this implies that missing values may not severely affect the SAM value.

Table 3
Overall classification accuracies of the LE dimensionality-reduced NDVI weekly time series. The mean and standard deviation (parentheses) of the overall classification accuracies are shown for 20 independent classifications; each time 0.5% of the study area CDL data locations were selected at random as training data and the remaining were used as test data.

Study area	LE-SAM accuracy	LE-SAM-R accuracy
Texas	56.4% (0.55%)	90.1% (0.20%)
Kansas	62.5% (0.40%)	78.2% (0.46%)
South Dakota	62.0% (0.53%)	77.2% (0.40%)

Table 4
Overall classification accuracies of the LE dimensionality-reduced 5-band reflectance weekly time series. The mean and standard deviation (parentheses) of the overall classification accuracies are shown for 20 independent classifications; each time 0.5% of the study area CDL data locations were selected at random as training data and the remaining were used as test data.

Study area	LE-SAM accuracy	LE-SAM-R accuracy
Texas	82.8% (0.35%)	90.3% (0.40%)
Kansas	83.2% (0.25%)	84.4% (0.26%)
South Dakota	79.3% (0.31%)	81.5% (0.44%)

4.2.2. LE-SAM

In the LE-SAM DR method, the SAM (Eq. 5) is used rather than the Euclidean distance (Eq. 1) to find the k nearest neighbors for every pixel in feature space so that missing data can be handled. In this study, k was selected as an empirical value of 40. If specific feature space dimensions values for either pixels a or b are missing, then the SAM is computed by ignoring the missing dimensions (the non-missing values in these dimensions are still used in the calculation of SAM for other pixels). The elements of the neighborhood graph G are computed not as (Eq. 2) but as:

$$g_{i,j} = SAM(a,b)^c \tag{6}$$

where $g_{i,i}$ is a defined element (i, j) in the matrix G, i and j are indices to pixels a and b, respectively, SAM(a, b) is defined as Eq. (5), and c controls the flatness of the kernel. In this study c was set to the empirical value of 2 suggested by (Yan & Niu, 2014); we found that similar results were obtained with $c = 1$ and $c = 3$, and this is likely because c only affects the edge weights in G that are unimportant compared with G's structures primarily determined by the selected distance metric.

Yan and Niu (2014) demonstrated that LE-SAM provided improved nonlinear DR of single date hyperspectral data over the conventional Euclidean distance based LE approach. However, their method did not have the refinement to handle missing data, and the hyperspectral bands containing missing data, which were saturated pixels, were simply discarded and not used in LE-SAM.

4.2.3. LE-SAM-R

Despite the implementation of the LE-SAM, the resulting DR bands will become less reliable when there are more missing data. Usually the patterns of missing satellite data are unevenly distributed in time (Brooks, Thomas, Wynne, & Coulston, 2012; Ju & Roy, 2008; Kovalskyy, Roy, Zhang, & Ju, 2011). The spatio-temporal dynamics of clouds at the times of satellite observation are not easy to predict. It is reasonable to assume, however, that cloud-free land surface observations are temporally correlated, i.e. the land surface state and condition observed on a given week will be similar to those observed a few weeks before or after (Ju, Roy, Shuai, & Schaaf, 2010). This property is used in the termed LE-SAM-R method to further reduce the sensitivity of LE-SAM to missing data that are present in satellite time series. Specifically, if the band values for either pixel a or b are missing on a certain week, then the missing values are replaced with the temporally closest non-missing pixel band values found by searching up to two weeks before or after the missing week. This process is described by Eqs. (7) and (8) defined for simplicity for a single band weekly time series over one year, i.e., for $n = 52$.

$$SAM_{refined}(a,b) = \frac{\sum_{week=1}^{52} x^a_{(week+shift_1(i\ \in\ 1\ldots9))} \cdot x^b_{(week+shift_2(i\ \in\ 1\ldots9))}}{\sqrt{\sum_{week=1}^{52}\left(x^a_{(week+shift_1(i\ \in\ 1\ldots9))}\right)^2}\sqrt{\sum_{week=1}^{52}\left(x^b_{(week+shift_2(i\ \in\ 1\ldots9))}\right)^2}} \tag{7}$$

$$\begin{aligned} shift_1 &\in \{0,\ -1,\ 0,\ 1,\ 0,\ -2,\ 0,\ 2,\ 0\} \\ shift_2 &\in \{0,\ 0,\ -1,\ 0,\ 1,\ 0,\ -2,\ 0,\ 2\} \end{aligned} \tag{8}$$

where a and b are two pixels whose feature space values are defined by $x_{z\in[1\ldots52]}$, and $shift_1$ and $shift_2$ are vectors that describe the possible temporal search window and are bounded in $[-2, 2]$ so that the maximum search range is ± 2 weeks. The shift vector values (Eq. 8) are tested sequentially with $i = 1\ldots9$ until $x^a_{(week+shift_1(i))}$ and $x^b_{(week+shift_2(i))}$ are both found to be valid observations, which are then used in the calculation of Eq. (7) for that week. If the band values for pixels a and b on a certain week are both missing, or there are no valid observations found for $i = 1\ldots9$, then $SAM_{refined}(a, b)$ is computed ignoring the band values for that week. The LE-SAM-R method is defined in the same way as LE-SAM but using $SAM_{refined}$ (Eq. 7) rather than SAM (Eq. 5), which will result in different k nearest neighbors for each pixel and different manifold structures represented by G.

The LE-SAM-R dimensionality reduction should be more reliable than LE-SAM when there are more missing observations in time series. The efficacy of LE-SAM-R will be reduced, however, if the land surface state or condition changes rapidly in periods of missing data. Therefore, a ± 2 week search window was used to increase the potential to replace missing band values with the temporally closest non-missing value over a persistent gap period of up to two weeks, while reducing the likelihood for surface change. A shorter duration ± 1 week search window was not used as persistent gaps greater than one week occurred quite frequently due to the cloudy nature of the study area time series data (Table 2). Longer duration search windows ($>\pm 2$ weeks) did not provide apparently improved or worse results than using the ± 2 week search window in the study areas. This is likely because the $SAM_{refined}$ (Eq. 7) used in LE-SAM-R with the ± 2 week search window is sufficiently reliable to capture the similarities among the Landsat weekly time series reflecting the phenology of various land cover types.

Table 5
Mean producer's and user's classification accuracies of the LE-SAM-R dimensionality-reduced 5-band reflectance weekly time series (corresponding to the overall accuracies in the rightmost column in Table 4). A total of 20 independent classifications were performed with 0.5% training data. Results are shown for major CDL classes covering >2% of each study area.

		Grass/pasture	Winter wheat	Corn	Cotton
Texas	Producer's accuracy	94.5%	84.4%	91.7%	49.1%
	User's accuracy	92.0%	90.9%	85.9%	69.7%

		Grass/pasture	Corn	Winter wheat	Fallow/idle cropland	Sorghum	Developed/open space	Alfalfa
Kansas	Producer's accuracy	92.9%	91.0%	82.1%	81.7%	80.4%	0.3%	71.6%
	User's accuracy	84.5%	87.1%	86.5%	81.3%	77.9%	8.3%	82.5%

		Grass/pasture	Winter wheat	Spring wheat	Corn	Sunflower	Water	Soybean	Hay	Developed/open space
South Dakota	Producer's accuracy	92.2%	82.2%	81.2%	77.1%	78.8%	93.3%	66.4%	6.1%	0.1%
	User's accuracy	81.9%	83.8%	82.9%	78.4%	78.4%	93.0%	69.5%	33.8%	2.7%

Table 6
Metrics overall classification accuracies. Mean and standard deviation (parentheses) of the overall classification accuracies are shown for 20 independent classifications; each time 0.5% of the study area CDL data locations were selected at random as training data and the remaining were used as test data.

Study area	Metrics accuracy
Texas	87.8% (0.46%)
Kansas	78.3% (0.42%)
South Dakota	77.7% (0.38%)

5. Analysis methodology

5.1. Qualitative Laplacian Eigenmaps feature extraction demonstration

The LE-SAM-R DR method was applied to 52 weeks of Landsat weekly reflectance bands, i.e. bands 2 (green, 0.53–0.61 μm), 3 (red: 0.63–0.69 μm), 4 (near-infrared: 0.78–0.90 μm), 5 (middle-infrared: 1.55–1.75 μm), and 7 (middle-infrared: 2.09–2.35 μm). The shortest wavelength Landsat ETM + band 1 (blue: 0.45–0.52 μm) was not used for its over sensitivity to atmospheric scattering (Ju et al., 2012; Roy, Qin, et al., 2014). From this time series with a dimension of 260 (5 reflectance bands × 52 weeks), up to 260 DR bands can be generated. Rather than illustrating them all, only the DR bands with high spatial correspondence with the CDL classes were examined. Such correspondences exist because the LE method is locality-preserving and so enables the information pertaining to different classes to be enhanced in separate DR bands (Yan & Niu, 2014) sometimes referred to as extracted features (Zhang et al., 2013).

5.2. Quantitative classification experiments

In each study area, a series of quantitative classification experiments were undertaken on the DR bands generated by the two LE methods (LE-SAM and LE-SAM) applied to the NDVI time series (Section 5.2.1) and the 5-band reflectance time series (Section 5.2.2), and the derived metrics bands (Section 5.2.3). The same classification methodology (Section 5.2.4) was adopted for the all three sets of experiments described below.

5.2.1. Classification of LE dimensionality-reduced NDVI time series

The DR bands generated by the LE-SAM and LE-SAM-R applied to 52 weeks of Landsat NDVI weekly time series in climate year 2010 were classified. Comparison of the classification results provides insights into the functionalities of these two LE DR techniques applied to a one-dimensional time series. The NDVI is examined as it has been widely used to study agriculture (Clerici, Weissteiner, & Gerard, 2012; Jakubauskas, Peterson, Kastens, & Legates, 2002; Ozdogan, 2010; Sakamoto et al., 2010; Wardlow & Egbert, 2008; Yan & Roy, 2014). Only the 20 DR bands associated with the 20 smallest non-zero eigenvalues were classified. This number of DR bands was selected to guarantee that sufficient information was maintained in the DR bands used for classification considering the maximum number of major CDL classes (nine classes in the South Dakota study area), while ensuring a relatively high dimensionality reduction ratio. It was not possible to apply the methods that have been developed to estimate a suitable minimum number of DR bands, as they cannot handle missing data (Hasanlou & Farhad, 2012; Levina & Bickel, 2004). We note, however, that using 20 DR bands provided a greater degree of data reduction than that provided by the 69 metrics bands (described in Section 5.2.3).

5.2.2. Classification of LE dimensionality-reduced multispectral reflectance time series

The DR bands generated by the LE-SAM and LE-SAM-R applied to 52 weeks of Landsat 5-band (bands 2, 3, 4, 5 and 7) reflectance weekly time series in climate year 2010 were classified. Comparison of the results with the classification results of the NDVI DR data (Section 5.2.1) provides insights into whether using the five reflectance bands provides improved classification results; this is expected as more spectral information are available to the dimensionality reduction. As for the NDVI time series experiments, only the 20 DR bands associated with the 20 smallest non-zero eigenvalues were classified.

5.2.3. Classification of metrics

The same five spectral bands (2, 3, 4, 5 and 7) and the same 52 weeks of Landsat data for climate year 2010 used for the LE-SAM and LE-SAM-R DR classification experiments were used to derive the metrics. The metrics were based on those used previously to classify 30 m percent tree cover, bare ground and other vegetation for all the conterminous United States (Hansen et al., 2011). Comparison of the classification results with the LE-SAM and LE-SAM-R DR classification results provides insights into whether using metrics or LE dimensionality-reduced data provides improved classification results.

In Hansen et al. (2011), temporal metrics that quantified the lowest, 2nd lowest, median, 2nd highest and highest values of bands 3, 4, 5, 7, and the NDVI, were extracted from the 2010 April to October WELD monthly products in an attempt to capture growing season phenology in a way that was robust to missing data. Hansen et al. (2011) also used metrics defined by Landsat reflectance bands 3, 4, 5, 7, the NDVI and simple band ratios (3/5, 3/7, 4/5, 4/7, 5/7), extracted from each of the 2010 WELD summer, autumn and annual composites. In this study, we added Landsat band 2 based metrics to ensure consistency with the LE classification experiments. Thus, we used the lowest, 2nd lowest, median, 2nd highest and highest values of bands 2, 3, 4, 5, 7, and the NDVI, extracted from the 2010 April to October WELD monthly products, providing 30 metrics. Further, we used Landsat reflectance bands 2, 3, 4, 5, 7, the NDVI and simple band ratios (2/5, 2/7, 3/5, 3/7, 4/5, 4/7, 5/7), extracted from each of the 2010 WELD summer, autumn and annual composites, providing a total of 39 metrics. In total 69 metrics were used. We note that this provides considerably more metric bands than the 20 DR bands.

Table 7
Mean producer's and user's classification accuracies of the metrics-based classifications (corresponding to the overall accuracies in Table 6). A total of 20 independent classifications were performed with 0.5% training data. Results are shown for major CDL classes covering >2% of each study area.

		Grass/pasture	Winter wheat	Corn	Cotton
Texas	Producer's accuracy	93.8%	76.8%	91.1%	39.4%
	User's accuracy	90.3%	83.1%	86.5%	65.0%

		Grass/pasture	Corn	Winter wheat	Fallow/idle cropland	Sorghum	Developed/open space	Alfalfa
Kansas	Producer's accuracy	89.9%	85.2%	66.3%	78.9%	72.8%	0.7%	68.7%
	User's accuracy	80.6%	84.6%	72.7%	74.1%	68.9%	18.3%	82.9%

		Grass/pasture	Winter wheat	Spring wheat	Corn	Sunflower	Water	Soybean	Hay	Developed/open space
South Dakota	Producer's accuracy	92.9%	72.6%	65.6%	65.4%	76.9%	92.9%	57.7%	8.3%	0.1%
	User's accuracy	78.9%	77.0%	71.2%	71.4%	75.3%	93.9%	71.6%	41.3%	5.2%

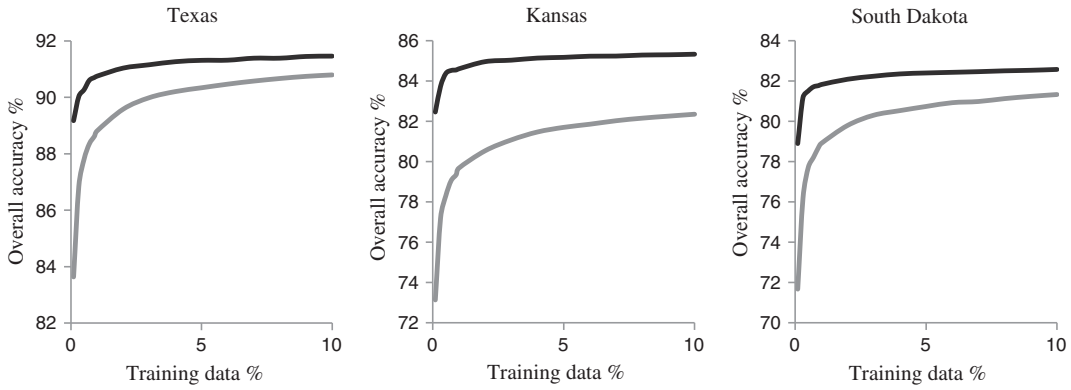

Fig. 6. Sensitivity of the overall classification accuracy to different training data sampling percentage (0.1%, 0.3%, 0.5%, 0.7%, 0.9%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9% and 10%). The mean overall accuracies of the LE-SAM-R dimensionality-reduced 5-band reflectance weekly time series (black) and the metrics (gray) are shown. The mean values are computed from a total of 20 independent classifications undertaken at each sampling percentage.

5.2.4. Classification experiment methodology

For the above classification experiments, the established supervised non-parametric random forest classifier was used. The random forest classifier can accommodate nonlinear relationships between variables and makes no assumptions concerning their statistical distributions (Breiman, 2001). It provides reduced likelihood of over-fitting training data by independently fitting a large number of decision trees, with each tree grown using a random subset of the training data and a limited number of randomly selected bands. The RANDOMFOREST software package (http://code.google.com/p/randomforest-matlab/) with default parameter settings was used — a total of 500 trees were grown with each tree considering 63.2% of the training data and considering four randomly selected predictor variables per tree.

Classification training and test data were extracted from the 2010 CDL data in the three study areas. Only the CDL classes covering over 2% of each study area were considered (i.e. the three, seven and nine CDL classes illustrated in Figs. 1–3 for the Texas, Kansas and South Dakota study areas, respectively). In the study areas, classification experiments were firstly undertaken using 0.5% of the CDL pixel locations selected at random to provide a total of 577, 609 and 525 training pixels for the Texas, Kansas and South Dakota study areas, respectively. The pixels of CDL classes covering ≤2% of each study area were excluded from the

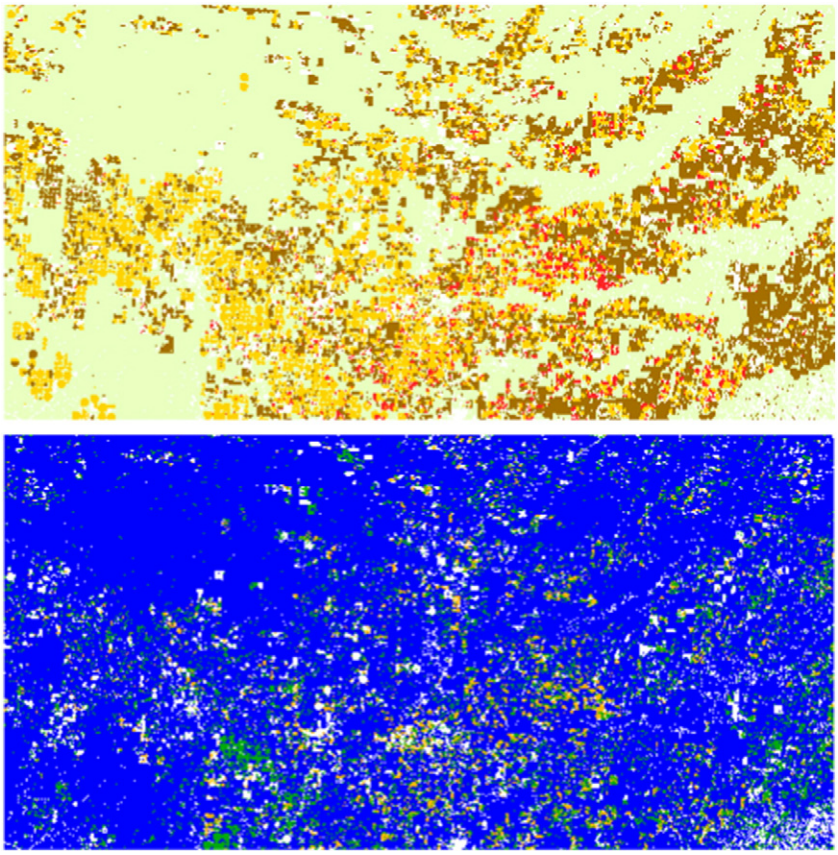

Fig. 7. Texas classification results. Top: Hard classification derived from 20 independent random forest classifications of the LE-SAM-R dimensionality-reduced 5-band reflectance weekly time series. Each classification was derived using 0.5% of the Texas study area CDL (Fig. 1) pixel locations as training data. Bottom: The "reliability" of the hard classification, showing the number of different classes (maximum 4) that the pixel was classified as over the 20 classifications: 1 (86.2%), 2 (10.1%), 3 (3.2%), 4 (0.5%). In both images, white shows the pixel locations where the CDL classes correspond to ≤2% of the study area (Fig. 1) and so were not classified.

training and test data selection. Sensitivity to the amount of training data was also examined by selecting at random 0.1%, 0.3%, 0.5% (already conducted), 0.7%, 0.9%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9% and 10% of the CDL pixels. Given a percentage of training data, the rest of the CDL pixels were used as test data. All of the classification experiments were repeated 20 times because of the relatively small percentages of training data used and because the CONUS 2010 CDL crop classification accuracy was imperfect.

The CDL class labels and the LE-SAM, LE-SAM-R and metrics band values were extracted at the same randomly selected training pixel locations. The training data were used to generate random forest classification results, which were then evaluated by the conventional classification accuracy statistics (overall, and per-class producer's and user's accuracies) derived from two-way confusion matrices using the corresponding test data (Congalton, 1991; Foody, 2002).

6. Results

6.1. Qualitative Laplacian Eigenmaps feature extraction demonstration

Fig. 5 (left) shows the four main Kansas CDL agricultural crops (corn, winter wheat, sorghum, and alfalfa crops) and (right) the four individual LE-SAM-R DR bands (linearly scaled to [0, 255] for illustration) that were observed visually to have the highest correspondence with these crops. The LE-SAM-R DR bands were obtained from the 52 weeks of Landsat 5-band reflectance weekly time series (described in Section 5.2.2). The Kansas study area was selected for illustration because it had an intermediate number of missing data (Table 2) and major CDL agricultural classes (four, Fig. 2). It is visually apparent that there is a high degree of consistency between the four CDL classes and the selected LE-SAM-R DR bands.

The LE-SAM-R DR bands for the other two study sites are not shown because of space constraints. For both the Texas and South Dakota study areas, the LE-SAM-R DR bands corresponding to the CDL crops covering more than 5% of each study area were easy to identify with high visual correspondence. However, the classes of cotton (3.3% of the study area) in Texas, developed/open space (3.4%) in Kansas, and hay (3.5%) and developed/open space (2.1%) in South Dakota had no clear corresponding LE-SAM-R DR bands — most likely because these crops occupied only a small fraction of the study areas. In addition, some other classes also had no clear corresponding LE-SAM-R DR bands, including grass/pasture (55.7%) in Texas, grass/pasture (35.2%) and fallow/idle cropland (11.6%) in Kansas, and grass/pasture (44.2%) in South Dakota. This is likely because these natural vegetation classes have less pronounced phenology than the agricultural classes (Fig. 4) and also because of their high within-class variability that does not allow for effective feature space partitioning (Georgescu, Shimshoni, & Meer, 2003; van der Maaten & Hinton, 2008).

6.2. Classification of LE dimensionality-reduced NDVI time series

The overall accuracies of the LE-SAM and LE-SAM-R dimensionality-reduced Landsat NDVI weekly time series classification results generated using 0.5% CDL training data are summarized in Table 3. Evidently, the LE-SAM-R approach provides consistently and unambiguously higher classification accuracies than LE-SAM. The >77% overall classification accuracies indicate that the LE-SAM-R method can effectively accommodate missing data present in the weekly NDVI time series, while the lower LE-SAM overall accuracies (<63%) indicate that it does not work as effectively.

6.3. Classification of LE dimensionality-reduced multispectral reflectance time series

The overall accuracies of the LE-SAM and LE-SAM-R dimensionality-reduced Landsat 5-band reflectance weekly time series classification results generated using 0.5% CDL training data are summarized in Table 4. For both of the LE methods, the overall classification accuracies are higher for the 5-band (Table 4) than the NDVI (Table 3) DR bands. In addition, the difference between the LE-SAM and LE-SAM-R is less apparent for the 5-band DR bands than the NDVI DR bands. This is because the LE methods have more spectral information when five bands rather than the NDVI (derived from the red and near-infrared bands) are used and this additional spectral information may compensate for any missing weekly gaps in the time series.

Table 4 shows that the LE-SAM-R overall classification accuracies are higher than those of LE-SAM for all three study areas. The greatest relative improvement of LE-SAM-R is for the Texas study area. As shown in Table 2, the Texas study area has the smallest percentage (30.3%) of valid weekly observations computed assuming 52 weeks of valid data, and the largest number of weeks (46 weeks) with at least one valid 30 m pixel across the study area. This suggests that the valid observations in the Texas time series data have the most uneven temporal distribution, which provide more opportunities for the missing-observation-adaptive

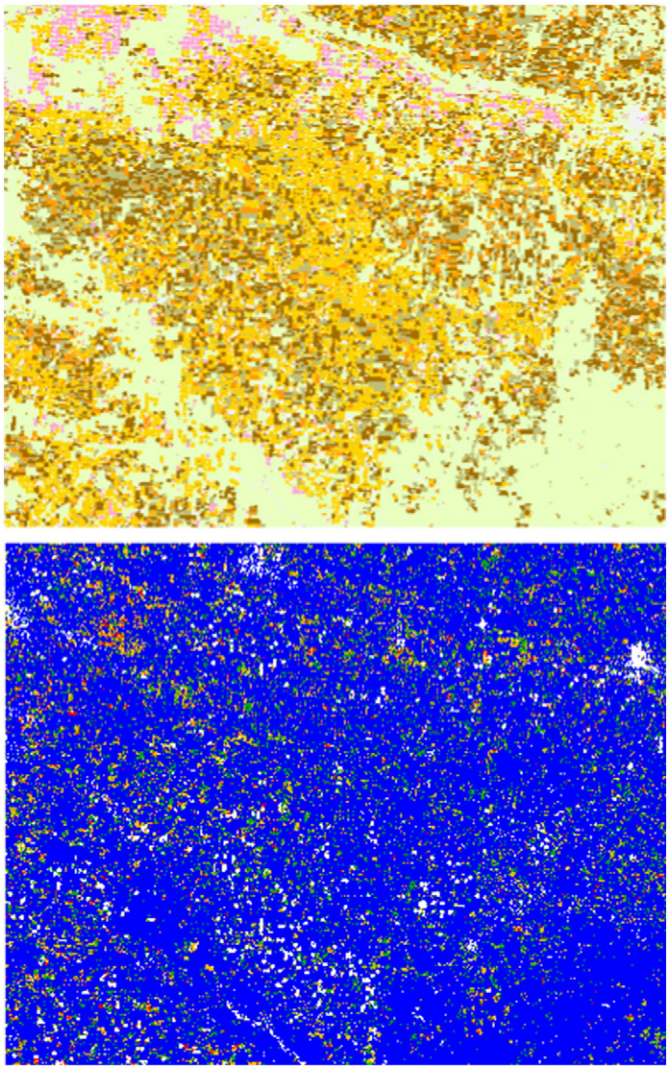

Fig. 8. Kansas classification results. Top: Hard classification derived from 20 independent random forest supervised classifications of the LE-SAM-R dimensionality-reduced 5-band reflectance weekly time series. Each classification was derived using 0.5% of the Kansas study area CDL (Fig. 2) pixel locations as training data. Bottom: The "reliability" of the hard classification, showing the number of different classes (maximum 7) that the pixel was classified as over the 20 classifications: 1 (83.4%), 2 (9.0%), 3–4 (6.5%), ≥5 (1.1%). In both images, white shows the pixel locations where the CDL classes correspond to ≤2% of the study area (Fig. 2) and so were not classified.

functionality provided by $SAM_{refined}$ (Eq. 7) (used by LE-SAM-R) to take effect than SAM (Eq. 5) (used by LE-SAM).

Table 5 summarizes the producer's and user's accuracies of the LE-SAM-R DR bands for each major crop class. The user's accuracy was calculated by dividing the number of all correctly classified pixels of a class by the sum of all pixels which had been assigned to that class; it indicates the probability that a pixel classified to a given class actually represents the reality on the ground (Congalton, 1991). The producer's accuracy was calculated by dividing the number of all correctly classified training pixels of a class by the sum of training data pixels for that class; it indicates the probability of a training pixel being correctly classified (Congalton, 1991). The producer's and user's accuracies were >77% for all the CDL classes covering more than 5% of each study area. The Kansas alfalfa class (3.1% of the study area) also had producer's and user's accuracies >71%. The Texas cotton class (3.3% of the study area) and the South Dakota soybean class (4.6% of the study area) had producer's and user's accuracies ranging from 49% to 70%. However, the South Dakota hay class (3.5% of the study area) had lower producer's (6.1%) and user's (33.8%) accuracies. The developed/open space class in the Kansas and South Dakota study areas also had classification accuracies <9% but it occupied less than 3.5% of each of the two study areas. These results indicate that reasonable per-class accuracies were obtained for the classes that occupy a non-minor portion of the images.

6.4. Classification of metrics

The overall, user's and producer's classification accuracies of the metrics-based classifications are summarized in Tables 6 and 7. The metrics overall classification accuracies (Table 6) are consistently and unambiguously lower than those provided by the LE-SAM-R dimensionality-reduced 5-band reflectance time series (right-most column in Table 4). The metrics overall classification accuracies are lower than those of the LE-SAM-R results by more than three standard deviations, specifically by 2.5%, 6.1% and 3.8% for the Texas, Kansas and South Dakota study areas, respectively.

Compared with the producer's and user's accuracies in Table 5, the corresponding accuracies reported in Table 7 are lower for the greater majority of the classes. When the producer's and user's accuracies are greater using the metrics, the difference is less than 1% for CDL classes covering more than 5% of each study area. These results indicate that the proposed LE-SAM-R method applied to 5-band reflectance weekly time series can provide overall higher classification accuracies than the conventional metrics approach.

6.5. Classification sensitivity to training data sampling

Classification accuracy is usually directly proportional to the training data set size, although careful training selection may help reduce the training set size without a loss of non-parametric supervised classification accuracy (Foody & Mathur, 2004). As DR techniques transform high-dimensional data to maximize the information content and minimize the noise (Bellman, 2003; Hughes, 1968), a more optimal dimensionality/data reduction method should enable a given degree of classification accuracy using fewer training data than a less optimal one.

Fig. 6 illustrates the overall classification accuracies provided by the LE-SAM-R applied to the Landsat 5-band reflectance weekly time series (black) and the overall accuracies provided by the metrics (gray), obtained using different percentages of training data. The smallest (0.1%) training percentage resulted in 115, 121 and 105 training pixels for the Texas, Kansas and South Dakota study areas respectively, which although small is an order of magnitude more than the number of CDL classes in the study areas. As before, a total of 20 independent classifications were performed at each training percentage. The overall classification accuracies illustrated in Fig. 6 are consistently higher for the LE-SAM-R approach than the metrics approach. Further, the LE-SAM-R overall classification accuracies are more stable than the metrics overall classification accuracies with fewer training data. The LE-SAM-R and the metrics overall classification accuracies become stable when approximately 0.5% and 2%, respectively, of the training data are used. The standard deviation of the overall classification accuracies of each set of 20 classifications are not shown, but for all the experiments they were less than 1%, except for the classifications undertaken using the least (0.1%) training data where the standard deviations were <2.3% for the three study areas. In all cases, the metrics

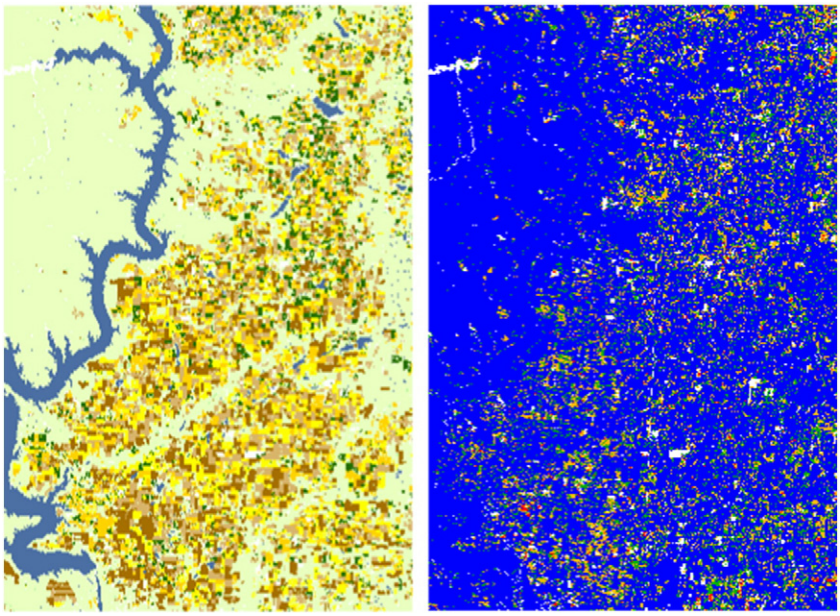

Fig. 9. South Dakota classification results. Left: Hard classification derived from 20 independent random forest supervised classifications of the LE-SAM-R dimensionality-reduced 5-band reflectance weekly time series. Each classification was derived using 0.5% of the South Dakota study area CDL (Fig. 3) pixel locations as training data. Bottom: The "reliability" of the hard classification, showing the number of different classes (maximum 9) that the pixel was classified as over the 20 classifications: 1 (81.4%), 2 (9.0%), 3–4 (8.9%), ≥5 (0.7%). In both images, white shows the pixel locations where the CDL classes correspond to ≤2% of the study area (Fig. 3) and so were not classified.

overall classification accuracies are lower than those of LE-SAM-R by more than three standard deviations. These results unambiguously indicate that the LE-SAM-R approach is more optimal than the metrics approach.

6.6. Spatially explicit LE-SAM-R hard classification results

This section illustrates the classification results and the classification "reliability" for the LE-SAM-R dimensionality-reduced Landsat 5-band reflectance weekly time series derived with a 0.5% training sample (described in Section 6.3). Rather than examine the 20 independent random classifications (the corresponding classification accuracy statistics are summarized in Tables 4 and 5), a single "hard" classification for each study area was generated for each study area. Classifications are described as "hard" when each pixel is classified into a single class category (Foody, 2000). Hard classification results were derived by allocating to each pixel the class that was most commonly occurring over the 20 random forest classifications (Figs. 7 and 8, top row; Fig. 9, left). Pixels where all 20 random forest classifications agreed are likely to be more reliable than those where there was disagreement (Dieye et al., 2012). To examine this, the number of different classes that each pixel was independently classified as over the 20 classifications was counted; the resulting reliability maps are also illustrated in Figs. 7 to 9.

The hard classification results have a high degree of visual consistency with the corresponding CDL data illustrated in Figs. 1–3, which reflects the high overall classification accuracies (90.3%, 84.4% and 81.5% for the Texas, Kansas and South Dakota study areas, respectively, Table 4). In addition, 86.2%, 83.4% and 81.4% of the Texas, Kansas and South Dakota study area pixels, respectively, were classified as only one unique class over the 20 independent classifications, which indicates the high reliability of the classifications. The least reliable pixels were found with the CDL classes covering less than 5% of the study areas, which were also the classes with low producer's and user's accuracies (Table 5). For example, the least reliable pixels in the Texas study area (classified into three or four classes over the 20 independent classifications, Fig. 7 bottom row) had high spatial correspondence with the CDL cotton class (Fig. 1), which is the class with the lowest producer's and user's accuracies (Table 5) in the Texas study area.

7. Summary and conclusions

Linear dimensionality reduction (DR) algorithms are well established and improved nonlinear DR algorithms have been demonstrated with hyperspectral data (Bachmann et al., 2005; Zhang et al., 2013). A major limitation of DR algorithms is that they cannot handle missing data (i.e., pixels that are unobserved or flagged as cloudy). If a nonlinear DR technique that is insensitive to missing data can be developed, then the possibility for improved satellite time series land cover classification and other information extraction approaches may be possible. In this paper, the Laplacian Eigenmaps (LE) nonlinear DR technique (Belkin & Niyogi, 2002) was refined for application to multispectral satellite time series that have missing data. The LE algorithm was refined (termed LE-SAM) using the spectral angle mapper (SAM) to find nearest neighbors for each pixel in feature space rather than using the conventional Euclidean distance. The SAM is insensitive to reflectance brightness variations and can be computed when values in some dimensions are missing. In addition, the LE-SAM algorithm was refined further (termed LE-SAM-R) by relaxing the constraint that the nearest neighbors for each pixel in feature space should be from the same time period, instead missing values could be replaced with the temporally closest non-missing pixel band values found by searching in adjacent temporal periods.

In the three agricultural study areas containing different amounts of missing data and land cover complexity, the utility of the two refined LE DR algorithms, which were applied to 52-week Landsat NDVI and multispectral reflectance time series, was demonstrated with a series of supervised random forest classification experiments using varying proportions of training data selected from USDA Cropland Data Layer (CDL) land cover maps. These classification results were compared with conventional random forest classifications using metrics derived from Landsat data (Hansen et al., 2011). Based on the experimental results, the following observations can be made.

i) The two refined LE DR methods, LE-SAM and LE-SAM-R, provided better classification performance when applied to the Landsat 5-band reflectance time series than the Landsat NDVI time series. This is as expected because the five reflectance bands convey more information than the NDVI (derived from just two reflectance bands). This underscores the utility of nonlinear DR approaches that are based on using multiple reflectance bands rather than a reduced set of spectral band ratios.

ii) The LE-SAM-R provided better classification performance than the LE-SAM for both the NDVI and reflectance time series. This indicates that the LE-SAM-R, which adopts a temporal search window in the calculation of SAM, is more effective than the LE-SAM in accommodating missing observations in the Landsat time series.

iii) The LE algorithm is locality-preserving (van der Maaten & Hinton, 2008) and so enables the information pertaining to different classes to be enhanced in separate DR bands (Yan & Niu, 2014). This was demonstrated by qualitative comparisons of the LE-SAM-R bands with CDL classes. Specifically, as demonstrated for the Kansas study area, the high consistency between the four major CDL crops and the four selected LE-SAM-R DR bands was evident. This indicates the potential of using single LE-SAM-R DR bands for preliminary mapping of single land cover classes, which is similar to the research demonstrated using the independent component analysis (ICA) linear DR algorithm applied to 8-day MODIS 500 m cloud-free time series (Ozdogan, 2010).

iv) Higher overall classification accuracies were obtained using the LE-SAM-R DR bands obtained from the Landsat 5-band reflectance weekly time series, compared with the conventional metrics-based approach, especially when smaller proportions of training data were used. When 0.5% of the study area CDL data were used for training, the metrics overall classification accuracies were lower than the LE-SAM-R results by 2.5%, 6.1% and 3.8% for the Texas, Kansas and South Dakota study areas, respectively, and the producer's and user's accuracies were lower for the greater majority of classes.

v) Compared with the conventional metrics-based approach, the number of training data required to achieve a given degree of classification accuracy was systematically lower using the LE-SAM-R DR bands obtained from the Landsat 5-band reflectance weekly time series.

Research to investigate the improved classification performance provided by LE-SAM-R, and in particular, experiments to ascertain which aspect of the algorithm (LE or $SAM_{refined}$) provides the most classification improvement when applied to multispectral reflectance time series, is recommended. The LE-SAM-R method is promising in that it is automated and provides dimensionality-reduced data that have desirable classification properties. These properties are important because land cover classification is labor intensive and far from automated; most large-area supervised classification projects spend the majority of their effort on training data collection (Townshend et al., 2012). It is well established that supervised classification accuracy is dependent on the quality and sampling of the training data (Breiman, 2001; Foody, 2002; Stehman, 2009). However, the optimal sampling needed to provide a given classification accuracy is nearly always unknown. As the dimensionality of the satellite data increases, e.g., with long

temporal data time series or hyperspectral data, the amount of training data required to capture the class variability reliably increases exponentially (Bellman, 2003; Hughes, 1968; Jimenez & Landgrebe, 1998). In practice, these issues are often ignored and instead are handled pragmatically by repeated training data collection and refinement, satellite data classification, and then classification accuracy assessment, until an acceptable classification accuracy is obtained.

Despite the desirable classification properties provided by the LE-SAM-R approach described in this study, we note that the computation required for the LE-SAM-R algorithm (and other manifold learning DR algorithms) increases geometrically with the image spatial dimensions (Bachmann et al., 2005, 2006; van der Maaten & Hinton, 2008). This is an issue, particularly for continental to global scale land cover classification (Friedl et al., 2010; Hansen et al., 2014). Future research to reduce the LE-SAM-R computational requirements, for example, by using the divide-conquer-and-merge strategies developed for isometric mapping (ISOMAP) nonlinear DR of hyperspectral data (Bachmann et al., 2005, 2006), approximation techniques developed for spectral-graph-based clustering applications (Mall, Langone, & Suykens, 2013; Yan et al., 2009), Graphics Processing Unit (GPU) enhanced computing (Bachmann, Ainsworth, Fusina, Topping, & Gates, 2010), and application of the LE-SAM-R independently to overlapping image subsets that are subsequently classified and then merged, will be investigated.

Information extraction from multi-temporal hyperspectral data is a research area that is only developing recently (Somers & Asner, 2013). Research to investigate the utility of the LE-SAM-R for application to multi-temporal hyperspectral data is recommended as fundamentally, hyperspectral data, like multispectral data, have missing data over time due primarily to cloud obscuration. This is especially pertinent as the availability of hyperspectral satellite data may increase with the launch of new hyperspectral satellite borne systems such as the Hyperspectral Infrared Imager (HysplRI) and the Environmental Mapping and Analysis Program (EnMAP) (Middleton et al., 2013; Roberts, Quattrochi, Hulley, Hook, & Green, 2012; Stuffler et al., 2007).

Acknowledgments

This research was funded by NASA grants NNX08AL93A and NNX13AJ24A. The U.S. Landsat project management and staff at the USGS National Center for Earth Resources Observation and Science (EROS), Sioux Falls, South Dakota, are thanked for provision of the Landsat ETM + data used to make the WELD products. We thank the two anonymous reviewers for their comments that improved this paper.

References

Bachmann, C.M., Ainsworth, T.L., & Fusina, R.A. (2005). Exploiting manifold geometry in hyperspectral imagery. *IEEE Transaction of Geoscience and Remote Sensing, 43*, 441–454.

Bachmann, C.M., Ainsworth, T.L., & Fusina, R.A. (2006). Improved manifold coordinate representations of large-scale hyperspectral scenes. *IEEE Transaction of Geoscience and Remote Sensing, 44*, 2786–2803.

Bachmann, C.M., Ainsworth, T.L., Fusina, R.A., Topping, R., & Gates, T. (2010). Manifold coordinate representations of hyperspectral imagery: improvements in algorithm performance and computational efficiency. *Proceeding of 2010 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 25–30 July 2010, Honolulu, HI* (pp. 4244–4247).

Belkin, M., & Niyogi, P. (2002). Laplacian Eigenmaps and spectral techniques for embedding and clustering. *Advances in Neural Information Processing Systems, 1*, 585–591.

Bellman, R.E. (2003). *Dynamic Programming*. Courier Dover Publications978-0-486-42809-3.

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*, 341–358.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*, 5–32.

Broich, M., Hansen, M.C., Potapov, P., Adusei, B., Lindquist, E., & Stehman, S.V. (2011). Time-series analysis of multi-resolution optical imagery for quantifying forest cover loss in Sumatra and Kalimantan, Indonesia. *International Journal of Applied Earth Observation and Geoinformation, 13*, 277–291.

Brooks, E.B., Thomas, V.A., Wynne, R.H., & Coulston, J.W. (2012). Fitting the multitemporal curve: A Fourier series approach to the missing data problem in remote sensing analysis. *IEEE Transaction of Geoscience and Remote Sensing, 50*, 3340–3353.

Busetto, L., Meroni, M., & Colombo, R. (2008). Combining medium and coarse spatial resolution satellite data to improve the estimation of sub-pixel NDVI time series. *Remote Sensing of Environment, 112*, 118–131.

Chang, J., Hansen, M.C., Pittman, K., Carroll, M., & DiMiceli, C. (2007). Corn and soybean mapping in the United States using MODN time-series data sets. *Agronomy Journal, 99*, 1654–1664.

Chung, F.R.K. (1996). *Spectral Graph Theory*. Providence, RI: American Mathematical Society.

Clerici, N., Weissteiner, C.J., & Gerard, F. (2012). Exploring the use of MODIS NDVI-based phenology indicators for classifying forest general habitat categories. *Remote Sensing, 4*, 1781–1803.

Collins, J.B., & Woodcock, C.E. (1996). An assessment of several linear change detection techniques for mapping forest mortality using multitemporal landsat TM data. *Remote Sensing of Environment, 56*, 66–77.

Congalton, R.G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*, 35–46.

DeFries, R., Hansen, M., & Townshend, J. (1995). Global discrimination of land cover types from metrics derived from AVHRR Pathfinder data. *Remote Sensing of Environment, 54*, 209–222.

Dieye, A.M., Roy, D.P., Hanan, N.P., Liu, S., Hansen, M., & Toure, A. (2012). Sensitivity analysis of the GEMS soil organic carbon model to land cover land use classification uncertainties under different climate scenarios in Senegal. *Biogeosciences, 9*, 631–648.

Feilhauer, H., Faude, U., & Schmidtlein, S. (2011). Combining Isomap ordination and imaging spectroscopy to map continuous floristic gradients in a heterogeneous landscape. *Remote Sensing of Environment, 115*, 2513–2524.

Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics, 7*, 179–188.

Foody, G. (2000). Estimation of sub-pixel land cover composition in the presence of untrained classes. *Computers & Geosciences, 26*, 469–478.

Foody, G. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment, 80*, 185–201.

Foody, G.M., & Mathur, A. (2004). Toward intelligent training of supervised image classifications: Directing training data acquisition for SVM classification. *Remote Sensing of Environment, 93*, 107–117.

Friedl, M.A., Sulla-Menashe, D., Tan, B., Schneider, A., Ramankutty, N., Sibley, A., et al. (2010). MODIS Collection 5 global land cover: Algorithm refinements and characterization of new datasets. *Remote Sensing of Environment, 114*, 168–182.

Friedman, J.H., & Tukey, J.W. (1974). A projection pursuit algorithm for exploratory data analysis. *IEEE Transactions on Computers, C-23*, 881–890.

Georgescu, B., Shimshoni, I., & Meer, P. (2003). Mean shift based clustering in high dimensions: A texture classification example. *Proceedings of 9th IEEE International Conference on Computer Vision, 13–16 October 2003, Nice, France* (pp. 456–463).

Green, A.A., Berman, M., Switzer, P., & Craig, M.D. (1988). A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Transaction of Geoscience and Remote Sensing, 26*, 65–74.

Hall, F.G., & Badhwar, G.D. (1987). Signature-extendable technology: Global space-based crop recognition. *IEEE Transaction of Geoscience and Remote Sensing, GE-25*, 93–103.

Han, T., & Goodenough, D.G. (2005). Nonlinear feature extraction of hyperspectral data based on locally linear embedding (LLE). *Proceedings of Geoscience and Remote Sensing Symposium, 2*, 1237–1240.

Hansen, M.C., Egorov, A., Potapov, P.V., Stehman, S.V., Tyukavina, A., Turubanova, S.A., et al. (2014). Monitoring conterminous United States (CONUS) land cover change with Web-Enabled Landsat Data (WELD). *Remote Sensing of Environment, 140*, 466–484.

Hansen, M.C., Egorov, A., Roy, D.P., Potapov, P., Ju, J., Turubanova, S., et al. (2011). Continuous fields of land cover for the conterminous United States using Landsat data: first results from the Web-Enabled Landsat Data (WELD) project. *Remote Sensing Letters, 2*, 279–288.

Hansen, M.C., & Loveland, T.R. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment, 122*, 66–74.

Hasanlou, M., & Farhad, S. (2012). Comparative study of intrinsic dimensionality estimation and dimension reduction techniques on hyperspectral images using k-nn classifier. *IEEE Geoscience and Remote Sensing Letters, 9*, 1046–1050.

Holben, B. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing, 7*, 1417–1434.

Hubert, M., Rousseeuw, P.J., & Verboven, S.A. (2002). Fast method for robust principal components with applications to chemometrics. *Chemometrics and Intelligent Laboratory Systems, 60*, 101–111.

Hughes, G. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory, 14*, 55–63.

Hyvärinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks, 10*, 626–634.

Hyvärinen, A., & Oja, E. (2000). Independent component analysis: Algorithms and applications. *Neural Networks, 13*, 411–430.

Jakubauskas, M.E., Peterson, D.L., Kastens, J.H., & Legates, D.R. (2002). Time series remote sensing of landscape–vegetation interactions in the southern Great Plains. *Photogrammetric Engineering and Remote Sensing, 68*, 1021–1030.

Jimenez, L.O., & Landgrebe, D.A. (1998). Supervised classification in high-dimensional space: Geometrical, statistical, and asymptotical properties of multivariate data. *IEEE Transactions of Systems, Man, and Cybernetics — Part C: Applications and Reviews, 28*, 39–54.

Jimenez, L.O., & Landgrebe, D.A. (1999). Hyperspectral data analysis and supervised feature reduction via projection pursuit. *IEEE Transactions on Geoscience and Remote Sensing, 37*, 2653–2667.

Johnson, D.M. (2013). A 2010 map estimate of annually tilled cropland within the conterminous United States. *Agricultural Systems*, *144*, 95–105.
Johnson, M.D. (2014). An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. *Remote Sensing of Environment*, *141*, 116–128.
Johnson, D.M., & Mueller, R. (2010). The 2009 cropland data layer. *Photogrammetry Engineering & Remote Sensing*, *76*, 1201–1205.
Journaux, L., Foucherot, I., & Gouton, P. (2006). Reduction of the number of spectral bands in Landsat images: A comparison of linear and nonlinear methods. *Optical Engineering*, *45* (067002–067002–12).
Ju, J., & Roy, D.P. (2008). The availability of cloud-free Landsat ETM+ data over the conterminous United States and globally. *Remote Sensing of Environment*, *112*, 1196–1211.
Ju, J., Roy, D.P., Shuai, Y., & Schaaf, C. (2010). Development of an approach for generation of temporally complete daily nadir MODIS reflectance time series. *Remote Sensing of Environment*, *114*, 1–20.
Ju, J., Roy, D.P., Vermote, E., Masek, J., & Kovalskyy, V. (2012). Continental-scale validation of MODIS-based and LEDAPS Landsat ETM+ atmospheric correction methods. *Remote Sensing of Environment*, *122*, 175–184.
Keshava, N. (2004). Distance metrics and band selection in hyperspectral processing with applications to material identification and spectral libraries. *IEEE Transaction of Geoscience and Remote Sensing*, *42*, 1552–1565.
Kovalskyy, V., & Roy, D.P. (2013). The global availability of Landsat 5 TM and Landsat 7 ETM+ land surface observations and implications for global 30 m Landsat data product generation. *Remote Sensing of Environment*, *130*, 280–293.
Kovalskyy, V., Roy, D.P., Zhang, X., & Ju, J. (2011). The suitability of multi-temporal Web-Enabled Landsat Data (WELD) NDVI for phenological monitoring — A comparison with flux tower and MODIS NDVI. *Remote Sensing Letters*, *3*, 325–334.
Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, J.P., et al. (1993). The spectral image processing system (SIPS) — Interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, *44*, 145–163.
Lee, D.S., Storey, J.C., Choate, M.J., & Hayes, R. (2004). Four years of Landsat-7 on-orbit geometric calibration and performance. *IEEE Transactions on Geoscience and Remote Sensing*, *42*, 2786–2795.
Levina, E., & Bickel, P.J. (2004). Maximum likelihood estimation of intrinsic dimension. *Proceedings of Neural Information Processing Systems, 13–18 December 2004, Vancouver, Canada*.
Loveland, T.R., & Dwyer, J.L. (2012). Landsat: Building a strong future. *Remote Sensing of Environment*, *122*, 22–29.
Mall, R., Langone, R., & Suykens, J.A.K. (2013). Kernel spectral clustering for big data networks. *Entropy*, *15*, 1567–1586.
Markham, B.L., Storey, J.C., Williams, D.L., & Irons, J.R. (2004). Landsat sensor performance: History and current status. *IEEE Transactions on Geoscience and Remote Sensing*, *42*, 2691–2694.
Martinez, A.M., & Kak, A.C. (2001). PCA versus LDA. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *23*, 228–233.
Middleton, E.M., Ungar, S.G., Mandl, D.J., Ong, L., Frye, S.W., Campbell, P.E., et al. (2013). The Earth Observing One (EO-1) Satellite Mission: Over a decade in space. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *6*, 243–256.
Murthy, C.S., Raju, P.V., & Badrinath, K.V.S. (2003). Classification of wheat crop with multi-temporal images: Performance of maximum likelihood and artificial neural networks. *International Journal of Remote Sensing*, *24*, 4871–4890.
Ozdogan, M. (2010). The spatial distribution of crop types from MODIS data: Temporal unmixing using independent component analysis. *Remote Sensing of Environment*, *114*, 1190–1204.
Rao, N.R. (2008). Development of a crop-specific spectral library and discrimination of various agricultural crop varieties using hyperspectral imagery. *International Journal of Remote Sensing*, *29*, 131–144.
Roberts, D.A., Quattrochi, D.A., Hulley, G.C., Hook, S.J., & Green, R.O. (2012). Synergies between VSWIR and TIR data for the urban environment: An evaluation of the potential for the Hyperspectral Infrared Imager (HyspIRI) Decadal Survey mission. *Remote Sensing of Environment*, *117*, 83–101.
Rogge, D., Bachmanna, M., Rivard, B., Nielsen, A.A., & Feng, J. (2014). A spatial–spectral approach for deriving high signal quality eigenvectors for remote sensing image transformations. *International Journal of Applied Earth Observation and Geoinformation*, *26*, 387–398.
Roy, D., Borak, J., Devadiga, S., Wolfe, R., Zheng, M., & Descloitres, J. (2002). The MODIS land product quality assessment approach. *Remote Sensing of Environment*, *83*, 62–76.
Roy, D.P., Ju, J., Kline, K., Scaramuzza, P.L., Kovalskyy, V., Hansen, M., et al. (2010). Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, *114*, 35–49.
Roy, D.P., Lewis, P., Schaaf, C., Devadiga, S., & Boschetti, L. (2006). The Global impact of cloud on the production of MODIS bi-directional reflectance model based composites for terrestrial monitoring. *IEEE Geoscience and Remote Sensing Letters*, *3*, 452–456.
Roy, D.P., Qin, Y., Kovalskyy, V., Vermote, E.F., Ju, J., Egorov, A., et al. (2014). Conterminous United States demonstration and characterization of MODIS-based Landsat ETM+ atmospheric correction. *Remote Sensing of Environment*, *140*, 433–449.
Roy, D.P., Wulder, M.A., Loveland, T.R., Woodcock, C.E., Allen, R.G., Anderson, M.C., et al. (2014). Landsat-8: Science and product vision for terrestrial global change research. *Remote Sensing of Environment*, *145*, 154–172.
Sakamoto, T., Wardlow, B.D., Gitelson, A.A., Verma, S.B., Suyker, A.E., & Arkebauer, T.J. (2010). A two-step filtering approach for detecting maize and soybean phenology with time-series MODIS data. *Remote Sensing of Environment*, *114*, 2146–2159.
Schaaf, C., Gao, F., Strahler, A., Lucht, W., Li, X., Tsang, T., et al. (2002). First operational BRDF, albedo and nadir reflectance products from MODIS. *Remote Sensing of Environment*, *83*, 135–148.
Settle, J., & Campbell, N. (1998). On the errors of two estimators of sub-pixel fractional cover when mixing is linear. *IEEE Transactions on Geoscience and Remote Sensing*, *36*, 163–170.
Shuai, Y., Masek, J.G., Gao, F., Schaaf, C.B., & He, T. (2014). An approach for the long-term 30-m land surface snow-free albedo retrieval from historic Landsat surface reflectance and MODIS-based a priori anisotropy knowledge. *Remote Sensing of Environment*, *152*, 467–479.
Small, C. (2012). Spatiotemporal dimensionality and time–space characterization of multitemporal imagery. *Remote Sensing of Environment*, *124*, 793–809.
Somers, B., & Asner, G.P. (2013). Invasive species mapping in Hawaiian rainforests using multi-temporal hyperion spaceborne imaging spectroscopy. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *6*, 351–359.
Somers, B., Cools, K., Delalieux, S., Stuckens, J., van der Zande, D., Verstraeten, W.W., et al. (2009). Nonlinear hyperspectral mixture analysis for tree cover estimates in orchards. *Remote Sensing of Environment*, *113*, 1183–1193.
Stehman, S.V. (2009). Sampling designs for accuracy assessment of land cover. *International Journal of Remote Sensing*, *30*, 5243–5272.
Stuffler, T., Kaufmann, H., Hofer, S., Förster, K. -P., Schreier, G., Müller, A., et al. (2007). The EnMAP hyperspectral imager — An advanced optical payload for future applications in Earth observation programmes. *Acta Astronautica*, *61*(1–6), 115–120.
Tenenbaum, J.B., de Silva, V., & Langford, J.C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, *290*, 2319–2323.
Townshend, J.R.G., Justice, C.O., & Kalb, V. (1987). Characterization and classification of South American land cover types using satellite data. *International Journal of Remote Sensing*, *8*, 1189–1207.
Townshend, J.R., Masek, J.G., Huang, C., Vermote, E.F., Gao, F., Channan, S., et al. (2012). Global characterization and monitoring of forest cover using Landsat data: Opportunities and challenges. *International Journal of Digital Earth*, *5*, 373–397.
van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning*, *9*, 2579–2605.
van der Maaten, L.J.P., Postma, E.O., & van den Herik, H.J. (2009). Dimensionality reduction: A comparative review, Tilburg University Technical Report, TiCC-TR 2009-005. URL: http://homepage.tudelft.nl/19j49/Matlab_Toolbox_for_Dimensionality_Reduction.html
Verstraete, M.M., & Pinty, B. (1996). Designing optimal spectral indexes for remote sensing applications. *IEEE Transaction of Geoscience and Remote Sensing*, *34*, 1254–1265.
Wardlow, B.D., & Egbert, S.L. (2008). Large-area crop mapping using time-series MODIS 250 m NDVI data: An assessment for the U.S. Central Great Plains. *Remote Sensing of Environment*, *112*, 1096–1116.
White, J.C., Wulder, M.A., Hobart, G.W., Luther, J.E., Hermosilla, T., Griffiths, P., et al. (2014). Pixel-based image compositing for large-area dense time series applications and science. *Canadian Journal of Remote Sensing*, *40*(3), 192–212.
Yan, D., Huang, L., & Jordan, M. (2009). *Fast approximate spectral clustering.* Technical Report No. UCB/EECS-2009-4. Berkeley: University of California (URL: http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-45.pdf).
Yan, L., & Niu, X. (2014). Spectral-angle-based Laplacian Eigenmaps for nonlinear dimensionality reduction of hyperspectral imagery. *Photogrammetric Engineering & Remote Sensing*, *80*, 11–23.
Yan, L., & Roy, D.P. (2014). Automated crop field extraction from multi-temporal Web Enabled Landsat Data. *Remote Sensing of Environment*, *144*, 42–64.
Zhang, X.Y., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., et al. (2003). Monitoring vegetation phenology using MODIS. *Remote Sensing of Environment*, *84*, 471–475.
Zhang, L.P., Zhang, L.F., Tao, D.C., & Huang, X. (2013). Tensor discriminative locality alignment for hyperspectral image spectral–spatial feature extraction. *IEEE Transactions on Geoscience and Remote Sensing*, *51*, 242–256.
Zhong, J., & Wang, R. (2006). Multi-temporal remote sensing change detection based on independent component analysis. *International Journal of Remote Sensing*, *27*, 2055–2061.