Landsat time series-based multiyear spectral angle clustering (MSAC) model to monitor the inter-annual leaf senescence of exotic saltcedar

Chunyuan Diao[a], Le Wang[b],*

[a] Department of Geography and Geographic Information Science, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[b] Department of Geography, University at Buffalo, the State University of New York, Buffalo, NY 14261, USA

ARTICLE INFO

ABSTRACT

In the western United States, the rapid expansion of exotic saltcedar along riparian corridors has drastically altered landscape structures and ecosystem functions. Monitoring the geographical distribution and spatio-temporal dynamics of this invasive species is essentially critical to conduct the systematic restoration of affected riparian ecosystems. Previous studies indicated that the leaf senescence stage is the optimal time window to remotely monitor saltcedar distributions. Yet due to climate variability and anthropogenic forcing, the timing of saltcedar leaf senescence varies over space and time. Given that the saltcedar leaf senescence stage only lasts for a short temporal window (i.e., three or four weeks), pinpointing the appropriate Landsat image across years and locations without the expert knowledge is challenging. Remotely sensed phenological time series analysis provides a practical means to locate the leaf senescence date on a per-pixel basis. However, affected by temporal revisit rates and cloud contamination, Landsat time series can only capture very limited temporal segments of vegetation phenology. The lack of Landsat imagery throughout the year makes the conventional time series analysis difficult. In this study, we developed a multiyear spectral angle clustering (MSAC) model to monitor the inter-annual leaf senescence of saltcedar with limited Landsat imagery on a per-pixel basis. The MSAC model leverages the Landsat images across years to temporally predict the fall phenology of plant species in a single year, and constructs the synthesized time series of spectral signatures to estimate critical phenological transition dates. Results indicated that the MSAC model could guide the construction of the composite Landsat surface reflectance image to accommodate spatial and inter-annual variation in the timing of plant leaf senescence. The phenology-guided composite image from the MSAC model achieved a greater saltcedar mapping accuracy than any single Landsat image in 2004. The proposed MSAC model provides new insights in estimating phenological transition dates of plant species with limited Landsat imagery. It opens up unique opportunities to guide the selection of representative remotely sensed imagery on a per-pixel basis for repetitive saltcedar mapping over wide geographical regions.

1. Introduction

Biological invasions constitute a component of global environmental change and pose significant threats to ecosystem functioning and biodiversity (Vitousek et al., 1996). Non-native saltcedar (*Tamarix* spp.) was introduced to the United States from Eurasia in the early 1900s. Since then, it has spread along riparian corridors at rates up to 20 km per year. Significant portions of riparian ecosystem structures have been transformed accordingly (Di Tomaso, 1998; Graf, 1978; Robinson, 1965). Saltcedar is now the third most dominant woody plant in the western U.S. riparian ecosystems (Friedman et al., 2005). The rapid expansion of saltcedar has been cited for displacing native species, reducing biodiversity, depleting river flows, and increasing fire

frequency (Di Tomaso, 1998; Glenn and Nagler, 2005; Zavaleta, 2000). Various remediation and restoration efforts have been made by local, state, and federal governments to control saltcedar. Among the control strategies, the biological control agent, named saltcedar leaf beetle (*Diorhabda elongate*), has been increasingly released to defoliate saltcedar in riparian areas (Dudley and Deloach, 2004; Hultine et al., 2010). Yet more recent studies revealed that water salvage from saltcedar removal may be much less than previously expected, and potential reductions of saltcedar may have unintended negative impacts on avian habitat qualities (Nagler et al., 2008; Van Riper et al., 2008). The negative role of saltcedar in ecosystem functions has thus been called for reevaluation (Stromberg et al., 2009). To comprehensively reevaluate the effects of saltcedar invasion and biocontrol defoliation,

repeated monitoring of saltcedar distribution at the regional scale is urgently needed (Nagler et al., 2011).

Remote sensing offers a practical means to monitor the spatial distribution of saltcedar. Fine spatial resolution airborne and spaceborne data, with the ability to predict the detailed plant information, have been explored extensively to map saltcedar distributions in various riparian ecosystems (Everitt and Deloach, 1990; Hamada et al., 2007; Narumalani et al., 2006; Pu et al., 2008; Wang and Zhang, 2014; Yang et al., 2013). These remotely sensed images were often acquired during the late fall and early winter period when saltcedar foliage color turns yellowish-orange to orangish-brown (Everitt and Deloach, 1990). Yet effective saltcedar management requires repetitive regional-scale monitoring of saltcedar distribution, which is still not feasible with fine resolution imagery due to its limited spatial coverage and expensive acquisition cost. In recent studies, the widely available moderate resolution Landsat surface reflectance imagery has also been explored for mapping the geographical distribution of saltcedar (Diao and Wang, 2016a, 2016b; Evangelista et al., 2009; Silván-Cárdenas and Wang, 2010; Wang et al., 2013). Diao and Wang (2016a) used the monthly Landsat time series of the year 2005 to map saltcedar along the Rio Grande River, and found that throughout the phenological cycle, the leaf senescence stage was the most crucial in spectrally distinguishing saltcedar from associated native vegetation (e.g., mesquite and willow). However, due to climate variability and anthropogenic forcing, the timing of saltcedar leaf senescence may vary across locations and years (Friedman et al., 2011). This spatial and inter-annual variation in the timing of saltcedar leaf senescence makes it challenging to pinpoint appropriate remotely sensed imagery, as saltcedar leaf senescence stage only lasts for three or four weeks.

Phenology-based time series analysis provides a solution to accommodate this spatio-temporal variation in vegetation. Time series of satellite-derived vegetation index (e.g., normalized difference vegetation index [NDVI]) has been commonly applied to estimate phenological stages of vegetation on a per-pixel basis. To reduce the effects of atmospheric interference and cloud contamination, smoothing algorithms (e.g., piecewise logistic function, asymmetric Gaussian function, and Harmonic analysis) are usually employed first to preprocess the time series (Beck et al., 2006; Diao and Wang, 2014; Jonsson and Eklundh, 2002). Critical phenological transition dates (e.g., onset of leaf senescence) can then be estimated from the smoothed curve of time series. A variety of estimation methods have been developed, including threshold-based method (Lloyd, 1990; White et al., 1997), inflection point method (Moulin et al., 1997; Zhang et al., 2003), and autoregressive moving average method (Reed et al., 1994). To further evaluate the status of leaf senescence (i.e., extent of leaf coloration), Zhang and Goldberg (2011) developed a temporally-normalized brownness index through combining a mixture model with the smoothed NDVI curve from time series of Moderate Resolution Imaging Spectroradiometer (MODIS). Ji and Wang (2016) used the MODIS End of Season-Time product (http://phenology.cr.usgs.gov) to estimate the saltcedar leaf peak coloration date by assuming that the timing of saltcedar peak coloration was linearly correlated with that of leaf drop. To date, phenology-based time series analysis has mostly been conducted by tracking the change of vegetation index, which is commonly used as an indicator of vegetation activity and is calculated based on specific bands of satellite imagery (e.g., red and near-infrared bands for NDVI). Yet the use of vegetation index may be limited by its sensitivity to soil background conditions, atmospheric aerosols, and saturation at densely vegetated areas. Hufkens et al. (2012) found that different vegetation indices (i.e., NDVI, EVI, and excess green index) can exert substantial influence on estimation results of phenological transition dates at four deciduous forest sites in the US. These varied forms of vegetation indices make the selection of any particular one difficult. As a wide range of spectral wavelengths can be used to estimate the vegetation status, tracking the collective change of spectral signature along the plant phenological trajectory may be more desirable.

However, time series of this expanded range of spectral information collected by satellites has seldom been reported in the current literature.

Advances in quality and availability of satellite time series have proven fruitful in tracking vegetation phenological dynamics in response to climate variability. The vast majority of previous work has used data from MODIS, Advanced Very High Resolution Radiometer (AVHRR), and SPOT Vegetation (VGT), owing to their fine temporal resolution and wide-area coverage (Morisette et al., 2006; Morisette et al., 2009). However, phenology captured by these coarse spatial resolution images often represents the leaf development of vegetation communities (e.g., including both native and non-native plants) (Zhang et al., 2004; Zhang and Goldberg, 2011). This mixed phenology makes the estimation of timing of saltcedar leaf senescence difficult, especially in spatially heterogeneous landscapes. The finer spatial resolution Landsat provides a more appropriate platform for capturing the plant phenology (Diao and Wang, 2016a; Kovalskyy et al., 2012; Zhong et al., 2014). Yet the 16 days' revisit rate and the influence of cloud cover often result in a dearth of imagery throughout the year, which makes the construction of Landsat time series challenging. Fisher et al. (2006) compressed the Landsat data from 1984 to 2002 into a single representative year in terms of the acquisition day of year (DOY) to estimate average phenological transition dates. The inter-annual phenological variation was then taken into account using a phase offset term. Yet the offset term was assumed to be uniform over the entire image and may not be sufficient to explain the phenological variation over space and time. To date, our ability to capture the spatial and inter-annual variation of plant phenology using the Landsat time series is still very limited.

The objective of this research was to develop a multiyear spectral angle clustering (MSAC) model to monitor the inter-annual leaf senescence of exotic saltcedar using limited Landsat images. The MSAC model was proposed in this study to construct the phenology-guided composite Landsat surface reflectance image, so as to accommodate the spatial and inter-annual variation in the timing of plant leaf senescence. The constructed composite Landsat surface reflectance image was then employed to monitor the distribution of exotic saltcedar along the Rio Grande River (near the town of Candelaria, Texas), which is one of the most infested areas in US.

2. Study site and data

2.1. Study site

The study site is located along the Rio Grande River, near the town of Candelaria, Texas (104.69° W, 30.14° N) (Fig. 1). The climate in the region is semi-arid to arid, and topography is characterized by small valleys and canyons, with elevation ranging from 700 m to 1150 m. The Rio Grande River originates from the San Juan Mountains of southern Colorado and flows southward to the Gulf of Mexico. It is one of the most regulated rivers in the southwestern US. As a result of dam construction and water regulation, the vegetation community along the river corridor has been transformed profoundly over the last century. In the study site, the native mesquite (*Prosopis* spp.) and willow (*Salix* spp.) on the riverbanks have been largely replaced by saltcedar, and the native cottonwood (*Populus* spp.) has been completely eliminated. The density of saltcedar varies from place to place, including both monotypic and mixed stands.

2.2. Image acquisition and pre-processing

2.2.1. Landsat data

The study site is covered by two overlapping Landsat scenes (path 31, row 39, and path 32, row 39). Due to climate variability, the timing of saltcedar leaf senescence may vary from year to year. To estimate the timing of inter-annual leaf senescence of saltcedar, constructing the
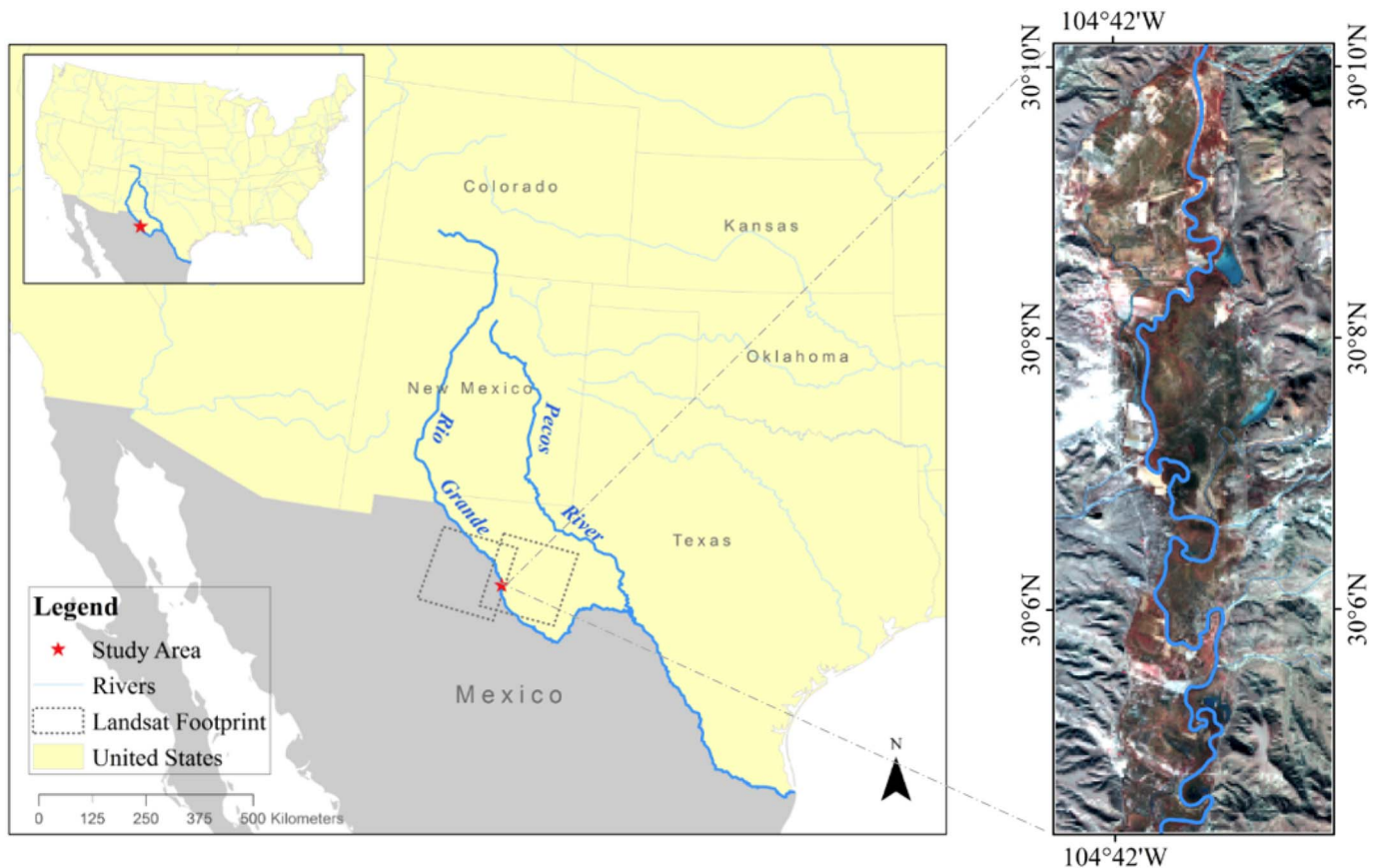
Fig. 1. Geographic location and false color Landsat imagery of the study site.

Landsat time series that can track its phenological trajectory is important. Affected by cloud cover and atmospheric conditions, Landsat images that are usable for saltcedar mapping can be very limited in certain years (e.g. 2004), which makes the building of conventional time series challenging. The proposed MSAC model will leverage the images in immediately preceding (e.g., 2003) and following (e.g., 2005) years to temporally predict the fall phenology in the image-lacking year (e.g., 2004), given the assumption that saltcedar fractions remain relatively stable during these three years. In this study, to construct the Landsat time series along the saltcedar leaf senescence trajectory in 2004, the cloud-free Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+) images spanning from August (peak greenness) to the following February (leaf-off) for the years 2003 to 2005 were obtained from the U.S. Geological Survey (USGS). The images were not evenly distributed during these three years, with 28 images in 2003, 15 images in 2004, and 25 images in 2005. The varying number of images each year also represents different levels of image availability in constructing the time series to estimate phenological characteristics. Due to the Scan Line Corrector (SLC) failure in 2003, Landsat ETM+ traces a zig-zag pattern along the satellite ground track and exhibits wedge-shaped data gaps. Thus the majority of Landsat ETM+ data over the three years contained gapped areas. As the devised MSAC model was constructed for each individual pixel to conduct temporal interpolation, the ETM+ pixels that were not in gapped areas were used. All the Landsat images were atmospherically corrected to the surface reflectance through the 6S radiative transfer model of Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) (Masek et al., 2006). The images were then georeferenced to the reference data (see Section 2.2.2) with root mean square errors less than 0.35. To investigate the collective role of a wide range of spectral wavelengths in tracing the phenological trajectory of saltcedar, six spectral bands (i.e., band 1, 2, 3, 4, 5, and 7) of Landsat images with the spatial resolution of 30 m were used in this study.

2.2.2. Reference data

The MSAC model was designed to temporally predict the fall phenology of saltcedar in 2004 to estimate the phenological transition date, which was then used to guide the construction of the composite image for mapping the saltcedar distribution (see Section 3). The performance of the composite image in classifying saltcedar was assessed by the reference data. Given the difficulty of obtaining sufficient ground reference data at the Landsat scale, the reference data were acquired from the Airborne Imaging Spectroradiometer for Applications (AISA) image. The AISA image was acquired on December 21, 2005, with the spatial resolution of 1 m and 61 bands in the spectral range from 400 to 1000 nm. Two field trips were conducted in November 2004 and December 2005 to collect the ground training and testing data for classifying the AISA image.

According to major land cover types in the study site, a two-level hierarchical classification scheme was designed (Table 1). Partitioning saltcedar into three classes at level II was necessary to avoid misclassification errors caused by its phenological variation. Spectral angle mapper (SAM) was utilized to classify the AISA image at level II, through comparing the spectral similarity between image spectra and reference spectra in the spectral library (Kruse et al., 1993). It assigned the image spectra to the class that formed the smallest spectral angle. The reference spectral library for training SAM was built from in-situ spectral signatures of typical land cover classes (e.g., saltcedar, native woody riparian species, and non-woody vegetation) measured by the ASD VNIR Field Spectrometer during the field trips, and image spectral signatures extracted directly for non-vegetated classes from the AISA image. The ASD Spectrometer was mounted on a boom 4 m above and perpendicular to the ground surface, to acquire canopy reflectance signatures of vegetated classes of the reference library. The level II classification result was aggregated to level I, the accuracy of which was assessed through the stratified sampling and estimation. For each

Table 1
Two-level hierarchical classification scheme of land covers in the study site.

Class level I	Class level II
Saltcedar	Green saltcedar (*Tamarix* spp.)
	Senescent saltcedar (*Tamarix* spp.)
	Leaf-off saltcedar (*Tamarix* spp.)
Non-saltcedar	Leaf-on willow (*Salix* spp.)
	Leaf-off willow (*Salix* spp.)
	Leaf-on mesquite (*Prosopis* spp.)
	Leaf-off mesquite (*Prosopis* spp.)
	Marshy weed (*Limnophila* spp.)
	Poverty weed (*Iva axillaris* Pursh)
	Green grasses
	Dry grasses
	Creosote bush (*Larrea tridentate*)
	Desert gravel
	Paved road
	Sand
	Roof
	Wetland
	Water (river, lake or pond)

Table 2
Confusion matrix of the AISA classification result.

		Reference data				
		Saltcedar	Non-saltcedar	Total	Map area (ha)	Proportion of mapped area
Classified data	Saltcedar	189	11	200	1086	0.42
	Non-saltcedar	5	195	200	1509	0.58
	Total	194	206	400	2595	1
	Overall accuracy = $189/200 * 0.42 + 195/200 * 0.58 = 96.24\%$					
	Saltcedar producer's accuracy = $(189/200 * 0.42) / (189/200 * 0.42 + 5/200 * 0.58) = 96.48\%$					
	Saltcedar user's accuracy = $189/200 = 94.50\%$					

mapped class (i.e., saltcedar or non-saltcedar), 200 pixels were randomly located and labelled according to the fieldwork and aerial photos. The area proportions of the mapped classes were the strata weights to obtain the stratified estimators (see Section 3.5 for stratified estimator). The stratified estimate of overall accuracy was 96.24%, and the producer's and user's accuracies of the saltcedar class were 96.48% and 94.50%, respectively (Table 2). This SAM classified image was further spatially resampled to the Landsat scale (30 m). The aggregated pixels served as the reference data to evaluate the proposed model, with fractions of saltcedar greater than 50% labelled as saltcedar (otherwise labelled as non-saltcedar). As thematic and spatial aggregations compensate for quantity and location errors, the aggregated reference data generated at the Landsat scale were assumed to be more accurate than those from the AISA classification result. Despite the great accuracy of the resampled Landsat-scale classified image, there may be uncertainty in the reference data, including sampling variability, class definitions, residual variability, and classification errors (Foody, 2010). In this study, we limited our attention to the uncertainty brought by sampling variability in the reference data, which was evaluated using a sensitivity analysis through randomly selecting training and testing samples 1000 times. Each time, according to the stratified sampling design, a total of 800 and 800 pixels were randomly selected from the resampled image for mapped saltcedar and non-saltcedar classes, respectively. Fifty percent of the sample for each class was employed to train the model, and the remainder was reserved for assessing the accuracy of classification results. Yet it is acknowledged that estimates of uncertainty using reference data would be slightly underestimated, as residual uncertainty was not addressed in this study.

3. Methods

3.1. Overview

A time series of spectral signatures was constructed for each individual pixel. The proposed MSAC model was designed on a per-pixel basis to estimate the phenological transition date using the images acquired across three years. The MSAC model mainly encompasses three components: time series spectral outlier removal, time series spectral clustering, and time series spectral matching (Fig. 2). For each pixel, only the spectral signatures with the NDVI values in the downward direction each year (i.e., DOY greater than that of maximum NDVI) were taken into account to monitor the inter-annual leaf senescence of saltcedar. Time series spectral outlier removal was devised to detect and remove the outliers in the spectral signatures of Landsat time series. In this study, Angle-Based Outlier Detection (ABOD) was used in time series spectral outlier removal to rule out outlying spectral signatures across years (Section 3.2). Time series spectral clustering was designed to cluster and organize the retained spectral signatures, as an appropriate sequence of multi-year spectral signatures along the temporal trajectory was critical to monitor inter-annual leaf senescence processes. In particular, *k*-means clustering algorithm was used in time series spectral clustering to organize and sort the multi-year spectral signatures (Section 3.3). With the sorted spectral signatures, time series spectral matching was devised to estimate the phenological transition date of leaf senescence. Spectral angle mapper-based moving average measure was developed in time series spectral matching to detect the spectral signature that exhibited the most remarkable change in the sorted time series. The corresponding date of that spectral signature was taken as the optimal phenological transition date (Section 3.4). In the devised MSAC model, angle of spectral signatures served as the measure of spectral similarity to reduce illumination and albedo effects. With the MSAC model, the composite Landsat image was then constructed. The spectral signature of each pixel in the composite image was acquired around the phenological transition date of leaf senescence. Finally, Random Forest was used to classify the composite image for mapping the distribution of saltcedar, and the accuracy was evaluated using the reference data (Section 3.5).

3.2. Time series spectral outlier removal

Time series spectral outlier removal was devised to remove the outliers in the spectral signatures of Landsat time series. Specifically, Angle-Based Outlier Detection (ABOD) was used in this study to preprocess the time series of spectral signatures across three years to reduce effects of atmospheric interference and cloud contamination (Kriegel et al., 2008; Pham and Pagh, 2012; Ye et al., 2014). The ABOD method measures the variance of angles between difference vectors of a data point to all other pairs in the time series, weighted by the distance of the points (Eq. (1)).

$$ABOD(A) = Var_{M,N \in D}\left(\frac{\overrightarrow{AM} \cdot \overrightarrow{AN}}{\|\overrightarrow{AM}\|^2 \cdot \|\overrightarrow{AN}\|^2} \right)$$
(1)

Here, A denotes the spectral signature of a pixel (i.e., 6 bands of Landsat) at a single date. D is the set of spectral signatures in the time series for the pixel. M and N are the spectral signatures of the pixel on any another two dates in the time series. A, M and N are called data points in the time series of spectral signatures D. $\overrightarrow{AM}$ and $\|\overrightarrow{AM}\|$ denote the difference vector $\overrightarrow{M} - \overrightarrow{A}$, and the distance between point A and point M, respectively.

The ABOD method calculates the variance of angles for each data point in the time series. A data point lying within a cluster of spectral signatures usually has a large value of variance in angles, as this data point is surrounded by other points in all directions. The angles between the difference vectors of this data point to all other pairs differ
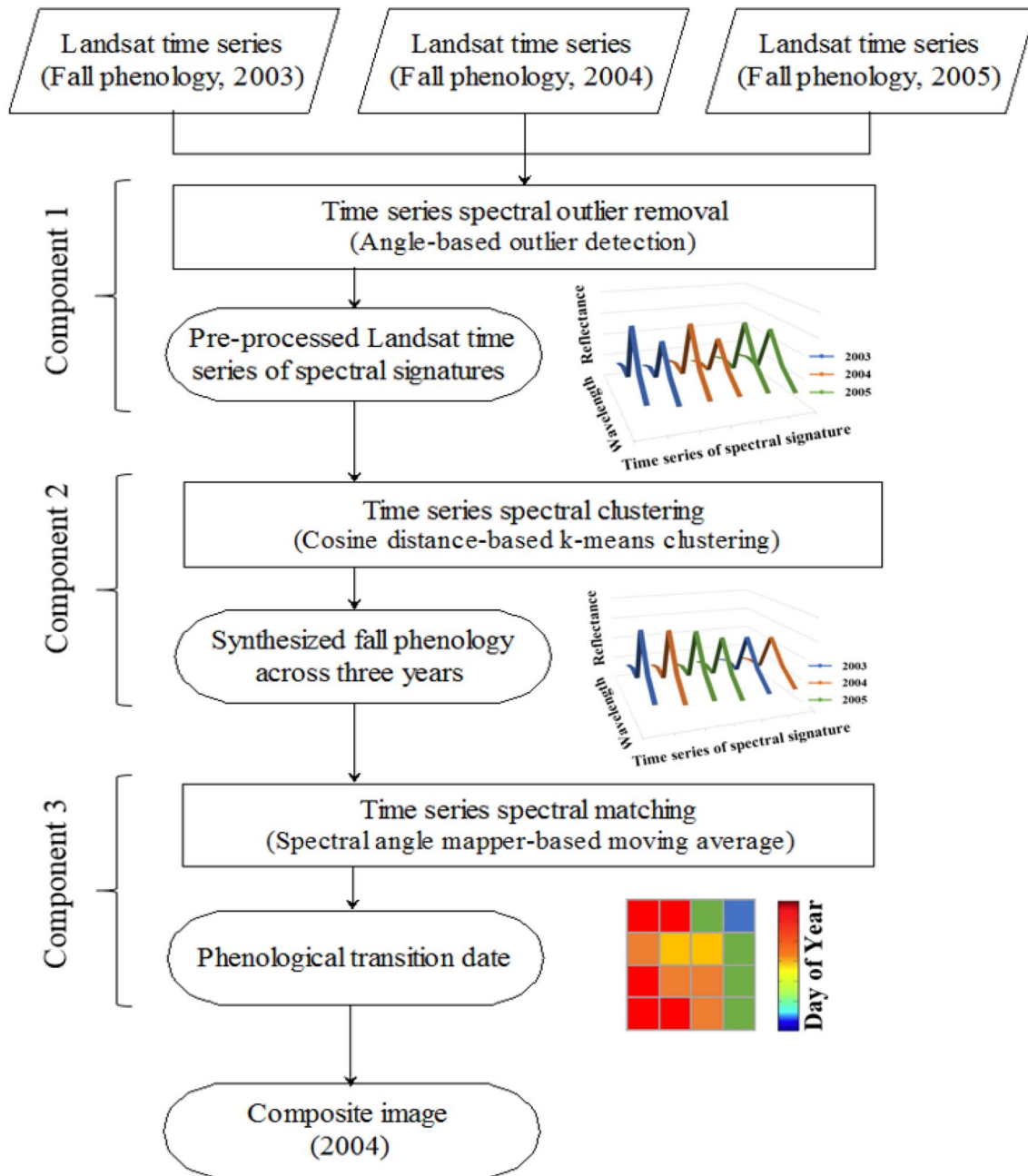
Fig. 2. Flowchart of the methodology.

widely. Yet an outlier typically has a small value of variance since the outlier is located in certain directions of most data points in the time series. The narrower the range of angles of a data point to other points is, the more likely it is an outlier. Therefore, the ABOD method can evaluate the degree of being an outlier of each data point in the time series.

Similar to the smoothing algorithms in preprocessing the NDVI time series, the ABOD method was used to remove the outliers in the Landsat time series of spectral signatures. The outliers typically had small ABOD scores and deviated markedly from other data points, which formed a gap between the scores of outliers and those of non-outliers. The gaps formed in the scores between potential outlying points and other data points were used to remove outliers. The ABOD method is parameter-free, and has been found to be more robust to an increasing dimensionality of data space than distance-based outlier detection methods (Kriegel et al., 2008). It can alleviate the effects of "curse of

dimensionality" and shows good performance with both low- and high-dimensional datasets.

3.3. Time series spectral clustering

After removing the outliers from the time series of spectral signatures, time series spectral clustering was designed to cluster and organize the retained spectral signatures. Unlike one-year time series of spectral signatures, the spectral signatures acquired across multiple years need to be organized to appropriately describe the change of spectral signatures along the leaf senescence trajectory (i.e., from leaf-on to leaf-off). In this study, the arrangement of multi-year spectral signatures took into account spectral similarity, NDVI, and DOY. The k-means algorithm was employed to group the multi-year spectral signatures into clusters (more information in following paragraphs) and the mean NDVI value of each cluster was calculated. The clusters were

sorted in terms of the descending order of mean NDVI values, since there was a decreasing trend of NDVI during the leaf senescence process of saltcedar. Within each cluster, the spectral signatures were sorted according to the descending order of NDVI, and this sequence was adjusted by the acquisition time of signatures each year, in which the sequence of spectral signatures can maintain a chronological order.

An appropriate sequence of multi-year spectral signatures along the temporal trajectory is crucial to monitor inter-annual leaf senescence processes. The k-means clustering algorithm was used in this study to organize the three-year spectral signatures in the downward direction of NDVI values and group similar spectral signatures into clusters (MacQueen, 1967). The k-means algorithm aims to partition the spectral signatures into k clusters, so as to minimize the within-cluster sum of squared distance. Instead of the Euclidean distance, the cosine distance was used in the k-means clustering algorithm in this study (Eq. (2)).

$$Dist(\mathbf{A}, \mathbf{C}) = 1 - \cos(\theta) = 1 - \frac{\overrightarrow{A} \cdot \overrightarrow{C}}{\|\overrightarrow{A}\|\|\overrightarrow{C}\|} \quad (2)$$

Here, A denotes the spectral signature of a pixel (i.e., 6 bands of Landsat) at a single date in the time series, and C denotes the centroid of a cluster of spectral signatures. $\|\overrightarrow{A}\|$ and $\|\overrightarrow{C}\|$ are the norms of the vectors $\overrightarrow{A}$ and $\overrightarrow{C}$, respectively. θ is the angle formed by $\overrightarrow{A}$ and $\overrightarrow{C}$. $Dist(\mathbf{A}, \mathbf{C})$ measures the similarity of orientation between $\overrightarrow{A}$ and $\overrightarrow{C}$, and is invariant to magnitude. Compared to the Euclidean distance, this angular-based distance measure can reduce the influence of illumination on the time series of spectral signatures, as variations in illumination mainly affect the length of vectors.

The k-means clustering algorithm requires the number of clusters (k) to be specified to identify the clusters of spectral signatures. The optimal number of clusters was estimated according to the Hartigan's method (Hartigan, 1975). As the number of clusters increases, the within-cluster sum of squared distance decreases. The optimal number of clusters was chosen when the reduction in the within-cluster sum of squares was not significant. As a heuristic algorithm, the convergence of the k-means clustering to a global minimum is heavily dependent on the initialization of centroids. In this study, the initial centroids were selected according to Arthur and Vassilvitskii (2007) to ensure that these centroid seeds were relatively far from each other. Besides, 100 replicates of random initialization of centroids were performed to locate the clusters of spectral signatures that could approximate the global minimum of the within-cluster sum of squared distance. Through time series spectral clustering and sorting, three years' spectral signatures were leveraged to track the spectral change of saltcedar along its leaf senescence trajectory.

3.4. Time series spectral matching

With the sorted multi-year spectral signatures, time series spectral matching was designed to estimate the phenological transition date of saltcedar leaf senescence. Two criteria were used in spectral matching to locate the desired transition date. First, following the temporal trajectory of leaf senescence (i.e., synthesized time series), the phenological transition date occurs when there is a noticeable change in the spectral signatures of saltcedar. It is expected that the spectral signature at the phenological transition date is appreciably different from preceding signatures in the synthesized time series. Second, the phenological transition date marks the onset of a new phenological process (i.e., leaf senescence). The spectral signature at the phenological transition date should be comparable to its immediate following spectral signatures.

Similar to the cosine distance in time series spectral clustering, spectral angle mapper (SAM) was used in time series spectral matching to compare the similarity between spectral signatures. SAM calculates the angle between two spectra in a multi-dimensional space, with

smaller angles indicating closer matches of the spectra. SAM has been found to be insensitive to albedo and illumination effects. To locate the phenological transition date, the spectral signature at each single date was compared to its previous spectral signatures in the synthesized time series using SAM (Eq. (3)). Following the second criterion, the spectral signature at each single date was also compared to its following spectral signatures using SAM (Eq. (4)).

$$MSACSM1(t) = \frac{\sum_{i=1}^{n_1} SAM(t, t-i)}{n_1} \quad (3)$$

Here, t denotes the position of a spectral signature in the sorted time series. n_1 is the number of its preceding spectral signatures to be compared. $SAM(t, t-i)$ stands for the spectral similarity between the spectral signature t and the spectral signature $t-i$ in the sorted time series measured by SAM. Each spectral signature in the sorted time series is compared to the immediate previous n_1 spectral signatures using SAM, and the average of SAM measures indicates how different this spectral signature is from the preceding signatures in the sorted time series.

$$MSACSM2(t) = \frac{\sum_{i=1}^{n_2} SAM(t, t+i)}{n_2} \quad (4)$$

Similar to Eq. (3), t denotes the position of a spectral signature in the sorted time series. n_2 is the number of its following spectral signatures to be compared. Each spectral signature in the sorted time series is compared to the immediate following n_2 spectral signatures using SAM, and the average of SAM measures indicates the extent of difference between this spectral signature and its following signatures.

The two moving average measures of SAM (i.e., Eqs. (3) and (4)) can be used to track the change of spectral signatures along the temporal trajectory of saltcedar leaf senescence. To meet both criteria, a spectral signature in the sorted time series needs to be compared to both its preceding and following signatures. The two moving average measures of SAM are thus integrated to be the time series spectral matching measure to estimate the phenological transition date (Eq. (5)).

$$MSACSM(t) = \frac{MSACSM1(t) + MSACSM2(t-1)}{2} \quad (5)$$

Here, $MSAC\_SM(t)$ is the spectral matching result of the spectral signature t in the sorted time series. At the phenological transition date, the spectral signature t will deviate from its previous signatures. Meanwhile, this spectral signature t will also be the onset of the following signatures in the time series, which implies that the spectral signature $t-1$ (i.e., the immediate previous signature) will differ from its following signatures. Hence a larger value of $MSAC\_SM(t)$ indicates that the date of the position t is more likely to be the phenological transition date.

With the time series spectral matching measure, the spectral signature that exhibited the most remarkable change in the sorted time series could be detected. The corresponding acquisition date of the signature was taken as the optimal phenological transition date across three years. The corresponding cluster of the signature was taken as the leaf senescence cluster, in which the spectral signatures were acquired during the leaf senescence stage, regardless of acquisition years. By this means, the images that were acquired during the leaf senescence stage could be located for each individual pixel.

The composite Landsat surface reflectance image of the year 2004, with each pixel acquired from its leaf senescence cluster, was then constructed. Among the images in the leaf senescence cluster, the image acquired in 2004 would be prioritized for the composite image. If two or more images from 2004 were in the cluster, the image lined close to that of the optimal phenological transition date in the sorted time series would be selected. For certain pixels, no image data were available during this period in 2004 due to cloud contamination. In that case the
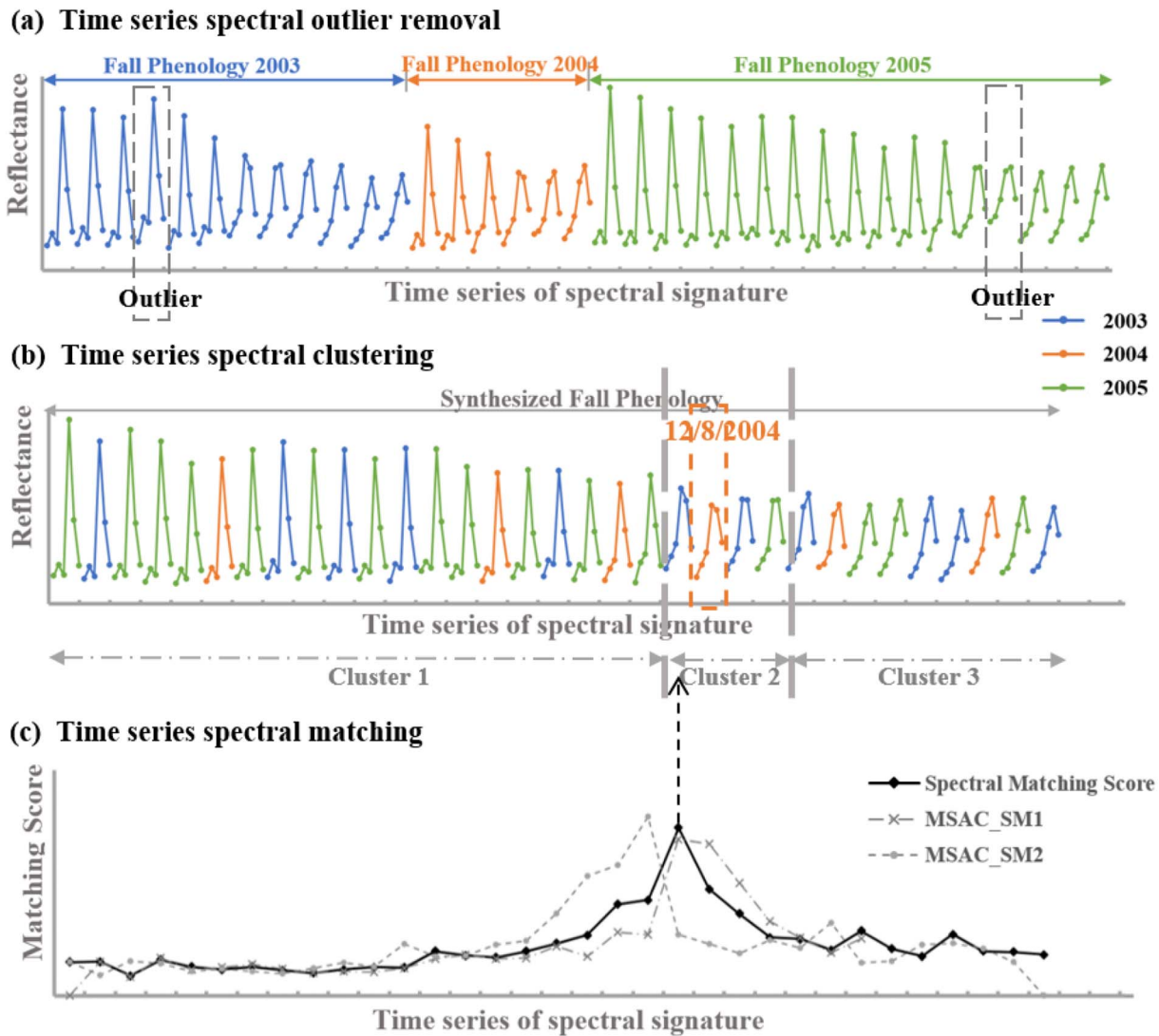
Fig. 3. Three components encompassed in the MSAC model to estimate the phenological transition date of saltcedar leaf senescence. Time series of spectral signatures are shown in two-dimensional space for the illustration purpose.

image acquired at the optimal phenological transition date would be selected. By temporally predicting the fall phenology of the year 2004 with the images across three years (i.e., 2004–2006), the MSAC model could guide the selection for the composite image, in which each pixel was ensured to be acquired during the leaf senescence period of saltcedar.

3.5. Accuracy assessment

The composite image was classified using the Random Forest (RF) algorithm to map the distribution of saltcedar in 2004 (Breiman, 2001). RF is a non-parametric ensemble learning method where a multitude of decision trees are constructed based on randomly selected features (i.e., bands) of a bootstrapped sample of training data. For each individual tree, a random subset of input features is selected at each tree node to determine a split, and the criterion for best splitting is based on the Gini index to maximize the dissimilarity between classes. The class of each pixel is then predicted by majority votes of all the decision trees in the forest, and the classification result is evaluated using the testing sample in terms of the confusion matrix. Given the stratified sampling design with saltcedar and non-saltcedar map classes used as two strata, the variation in estimation weights (e.g., different sampling intensities in different strata) associated with each stratum should be incorporated in

constructing the confusion matrix (Olofsson et al., 2014; Olofsson et al., 2013). Stratified estimators were used for which strata weights were the area proportions of the mapped classes. On this basis, the overall accuracy, producer's and user's accuracies of saltcedar were calculated using the reference data (i.e., testing sample with 400 and 400 pixels randomly selected for mapped saltcedar and non-saltcedar classes, respectively) to evaluate the performance of the composite image. The stratified estimator of area invaded by saltcedar was then produced from the composite image in conjunction with the reference data. A 95% confidence interval of the area of saltcedar invasion was provided to quantify the uncertainty of the estimated area obtained. To further assess the role of the composite image generated by the MSAC model in mapping saltcedar, the classification result of the composite image was compared to the largest classification accuracy achieved by a single Landsat image (i.e., the Landsat image acquired on December 8) in 2004.

The performance of the MSAC-based composite image was also compared to that of the conventional vegetation index-based composite image. The dearth of images in 2004 makes it difficult to conduct the conventional vegetation index-based time series analysis (e.g., fit a smoothing NDVI time series curve). Without fitting the time series curve, threshold-based estimation methods may be possible to estimate the phenological transition dates and construct the corresponding

composite image. As a threshold-based estimation method, the Midpoint method (i.e., the timing of NDVI reaching the half of amplitude) was used in this study to estimate the phenological transition date of leaf senescence. For each pixel, the spectral signature of the image acquired around the timing of NDVI decreasing to the half of its annual amplitude in the downward direction of phenological trajectory was used to construct the composite image. The NDVI-based composite image generated from the Midpoint threshold-based estimation method was also classified using RF and evaluated with the same reference data.

Additionally, the MSAC-based composite image was also evaluated according to the purity level of saltcedar within a pixel. The fractional coverage of saltcedar in a pixel affected the shape of its spectral signature, and the MSAC model was constructed based on the time series of spectral signatures. Thus the purity level of the pixel might affect the efficacy of the proposed model in estimating the saltcedar leaf senescence timing. The classification accuracies of the composite image for multiple pixel purity levels (i.e., 50% to 90% with an increment of 5%) were calculated. For example, the pixel purity level of 50% indicated that both training and testing sampling units were selected from the pixels with fractional coverage of saltcedar (or non-saltcedar) greater than 50%. The sample sizes for evaluating all the purity levels were the same. For each level, 800 pixels were randomly selected for training and the other 800 pixels were reserved for testing.

4. Results

4.1. Estimation of saltcedar leaf senescence timing

In this study, the timing of saltcedar leaf senescence in 2004 was estimated by leveraging all available Landsat images in the fall phenology of the years 2003, 2004, and 2005 using the proposed MSAC model. As an example, the time series of spectral signatures for a typical pure saltcedar pixel was shown in Fig. 3. The spectral signatures acquired in different years were shown in varied colors (12 images from 2003, 6 images from 2004, and 17 images from 2005).

Time series spectral outlier removal was used to rule out outlying spectral signatures that were affected by atmospheric and cloud interference (see Section 3.2). Specifically, two spectral signatures (one in 2003 and the other in 2005) were detected as outliers and removed due to small variances calculated by the ABOD method. The remaining spectral signatures were then clustered and organized using time series spectral clustering (see Section 3.3). Three clusters were formed according to the Hartigan's method. The clusters of spectral signatures across three years were sorted based on the mean NDVI value, and the spectral signatures within a cluster were then organized in terms of NDVI and DOY. The sorted time series of spectral signatures was shown in Fig. 3. The trend of the change in the spectral signatures along the time series was consonant with the phenological process of saltcedar leaf senescence (i.e., from leaf-on to leaf-off). This change in the spectral signatures was associated with the change in leaf color, leaf cell structure, and leaf water content of saltcedar. The spectral-based MSAC model can comprehensively consider these effects.

With the sorted multi-year spectral signatures, the timing of saltcedar leaf senescence was then estimated using time series spectral matching (see Section 3.4). For each spectral signature, multiple preceding and following spectral signatures (i.e., n1 and n2) were considered in time series spectral matching. Selecting the number of spectral signatures (i.e., moving average interval) to be compared is a critical issue, as a small interval may grasp insignificant trend changes and a large interval is less sensitive to the transition of phenological events. We experimented with several moving average intervals ranging from 2 to 10, and evaluated time series spectral matching results of pure saltcedar pixels. Among the moving average intervals, we found that a moving average interval of four spectral signatures achieved the largest spectral matching score at the phenological transition date of

saltcedar leaf senescence. Compared to other moving average intervals, it was more robust to insignificant trend changes while capturing the key phenological transition events. Thus a moving average interval of four spectral signatures was selected.

The time series spectral matching result for this saltcedar pixel was shown in Fig. 3. Along the time series, the spectral signature 21 attained the largest value of MSAC_SM1 and was perceived as the most different from its preceding four signatures. The spectral signature 20 had the largest value of MSAC_SM2 and varied the most from the following four signatures. By integrating the MSAC_SM1 and MSAC_SM2 measures, the spectral signature 21 obtained the largest spectral matching score and the corresponding date was selected as the optimal phenological transition date across three years. Given the similarity of spectral signatures within a cluster, the corresponding cluster was indicative of the leaf senescence stage of saltcedar. It was thus taken as the saltcedar leaf senescence cluster. Four spectral signatures were included in this cluster, with two acquired in 2003 (November 28 and December 6), one in 2004 (December 8), and one in 2005 (December 26). Thus in 2004, the most representative image to map the senescent saltcedar for this pixel was the one acquired on December 8.

4.2. Classification result of the composite image

The MSAC model guided the construction of the composite image, in which the spectral signature of each individual pixel was selected from the leaf senescence stage. Due to the scarcity of Landsat images in 2004, there may be no image data available during the leaf senescence stage in 2004 for some pixels. In this case, the spectral signature acquired at the optimal phenological transition date (i.e., attaining the largest spectral matching score) would be used to construct the composite image (Fig. 4). The varied color in the composite image acquisition dates indicated the spatial heterogeneity and variation in saltcedar leaf senescence timing. Four pixels in the composite image were highlighted, with two pixels acquired from the images in 2004 (i.e., November 29 and December 8), one pixel from the image in 2003 (November 28), and one pixel from the image in 2005 (December 26). Among the mapped area of 2595 ha in Fig. 4, pixels of 1607 ha were acquired from the images in 2004, pixels of 520 ha were from the images in 2003, and the rest of pixels of 468 ha were from the images in 2005. Hence the composite image can accommodate the spatial and inter-annual variation in saltcedar leaf senescence, and facilitate the repetitive regional-scale mapping of this invasive species.

The performance of the composite image in mapping the distribution of saltcedar was evaluated, and the confusion matrix that accounted for different sampling intensities in different strata was shown in Table 3. To account for the stratified sampling design, stratified estimators were used for which strata weights were equal to the area proportions of the mapped types (see Table 3 for details). The overall classification accuracy of the composite image was 81.26%. The corresponding producer's and user's accuracies of the saltcedar class were 70.14% and 83.50%, respectively. To evaluate the uncertainty of sampling variability in the reference data (resampled AISA classification result), the training and testing samples were randomly selected 1000 times. Each time the classification accuracy of the composite image was assessed with the corresponding testing sample using the stratified estimates. The mean overall accuracy (95% confidence interval), mean producer's accuracy of saltcedar (95% confidence interval), and mean user's accuracy of saltcedar (95% confidence interval) were 80.35% (79.34%–81.97%), 82.69% (81.67%–83.93%), and 69.27% (68.23%–70.98%), respectively. Given the similar evaluation accuracy from 1000 simulations and great classification accuracy of the AISA imagery, we randomly selected one sample as the reference data to conduct the following evaluation. To further assess the role of the composite image generated by the MSAC model in mapping saltcedar, the classification result of the composite image was compared to the largest classification accuracy achieved by a single Landsat image in
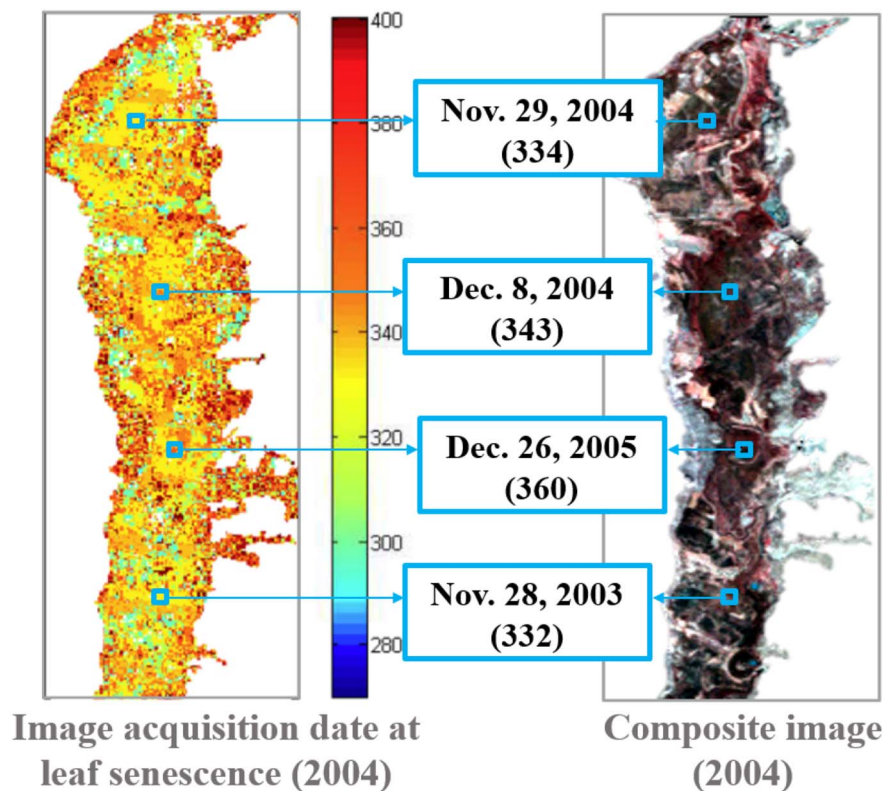
Fig. 4. Composite image in 2004 generated by the MSAC model.

2004. Among all the Landsat images in 2004, the image acquired on December 8 achieved the greatest performance, with the stratified estimate of overall accuracy being 74.79%. The producer's and user's accuracies of the saltcedar class using this single-date Landsat image were 61.22% and 80.25%, respectively.

Compared to the accuracies of the singe-date image on December 8, the user's accuracy of the saltcedar class increased by about 3% and the producer's accuracy of the saltcedar class increased by 9% through using the composite image. The improved accuracy could also be reflected in the confusion matrices constructed for these two classification strategies (Table 3). The comparisons revealed that the areas saltcedar occupied were underestimated with the traditional single-image-based mapping strategy, mainly due to the spatial variation in the timing of saltcedar leaf senescence. Saltcedar in various phenological stages (e.g., leaf on, leaf senescence, and leaf off) may exist within an image, yet the single-date image is only capable of mapping saltcedar during its senescent stage. Saltcedar at other phenological stages (e.g., leaf-on and leaf-off) are more inclined to be confused with native vegetation. The composite image was constructed on a per-pixel basis, with the spectral signature of each pixel acquired during the saltcedar leaf senescence stage. It was beneficial to correct the confusion errors by accommodating the spatial and inter-annual phenological variation.

Pinpointing the appropriate image for remotely mapping saltcedar is crucial, as classification results using different single-date Landsat images can vary widely. Table 3 shows the classification accuracies of two single-date Landsat images, acquired on November 29, 2004 and December 8, 2004, respectively. Compared to the image on December 8, the classification accuracy of the image acquired on November 29 deteriorated appreciably. The stratified estimate of overall accuracy dropped from 74.79% to 71.62%, and the producer's accuracy of saltcedar decreased from 61.22% to 56.60%. As saltcedar leaf senescence stage only lasts for a short temporal window (i.e., three or four weeks), one-week difference in image acquisition date can exert considerable influence on capturing colored saltcedar. The composite Landsat image constructed by the proposed model can not only accommodate this

phenological variation, but also ease the challenge to select the appropriate Landsat image repeatedly over space and time in saltcedar mapping.

The performance of the MSAC-based composite image was also compared to that of the NDVI-based composite image (Table 4). The NDVI-based composite image was constructed with the spectral signature of each pixel acquired around the timing of NDVI dropping to the half of amplitude. Compared to the MSAC-based composite image, the NDVI-based composite image yielded a much smaller mapping accuracy. The stratified estimate of overall accuracy of the NDVI-based composite image was 73.66%. The producer's and user's accuracies of saltcedar class were 59.49% and 79.00%, respectively. This noticeable decrease in the performance can mainly be explained by two reasons. On one hand, the leaf senescence timing in the NDVI-based composite image is defined as the timing of NDVI dropping to the half of amplitude. However, it is still not clear how the half of NDVI amplitude corresponds to the status of leaf coloration, as the threshold value (e.g., half of amplitude) to monitor the leaf senescence timing also varies across locations and species. Even more challenging is that the dearth of images in 2004 impedes the conventional time series analysis (e.g., smoothing). The difficulty of pre-processing the NDVI values in the fall phenology makes the construction of the NDVI-based composite image sensitive to the influence of atmospheric noise. As a result, the overall accuracy of NDVI-based composite image was less than that of the single Landsat image acquired on December 8 (Table 3). On the other hand, there may be no image data available during the leaf senescence stage for some pixels, due to the lack of images in 2004. Unlike the NDVI-based method, the MSAC model can take advantage of the Landsat images from multiple years to compensate the lack of images in a single year. It synthesizes the phenological information across years to temporally predict the fall phenology in 2004. The MSAC-based composite image is more likely to capture the senescent saltcedar, and thus achieves a larger mapping accuracy.

The classification map of the MSAC-based composite image in 2004 was shown in Fig. 5. The areas invaded by saltcedar were shown in

Table 3

Confusion matrices for the MSAC-based composite Landsat image and two single Landsat images.

MSAC-based composite image (2004)

		Reference Data				
		Saltcedar	Non-saltcedar	Total	Map area (ha)	Proportion of mapped area
Classified data	Saltcedar	334	66	400	946	0.36
	Non-saltcedar	80	320	400	1649	0.64
	Total	414	386	800	2595	1
	Overall accuracy = 334/400*0.36 + 320/400*0.64 = 81.26%					
	Saltcedar producer's accuracy = (334/400*0.36)/ (334/400*0.36 + 80/400*0.64) = 70.14%					
	Saltcedar user's accuracy = 334/400 = 83.50%					

Single Landsat image (Dec. 8, 2004)

		Reference Data				
		Saltcedar	Non-saltcedar	Total	Map area (ha)	Proportion of mapped area
Classified data	Saltcedar	321	79	400	927	0.36
	Non-saltcedar	113	287	400	1668	0.64
	Total	434	366	800	2595	1
	Overall accuracy = 74.79%					
	Saltcedar producer's accuracy = 61.22%					
	Saltcedar user's accuracy = 80.25%					

Single Landsat image (Nov. 29, 2004)

		Reference Data				
		Saltcedar	Non-saltcedar	Total	Map area (ha)	Proportion of mapped area
Classified data	Saltcedar	310	90	400	899	0.35
	Non-saltcedar	126	274	400	1696	0.65
	Total	436	364	800	2595	1
	Overall accuracy = 71.62%					
	Saltcedar producer's accuracy = 56.60%					
	Saltcedar user's accuracy = 77.50%					

Note: The accuracies of MSAC-based composite image and two Landsat images were evaluated using stratified estimators.

yellow. Overall, a great portion of the study area was susceptible to saltcedar invasion and highly dense areas were mostly distributed along the river. The stratified estimate of the area invaded by saltcedar (± 95% confidence interval) was 1112 ha ± 75 ha. In the study site, the expansion of saltcedar along the riparian corridor posed serious threats to the native vegetation communities and biodiversity. The native mesquite and willow on the riverbanks have been largely replaced by saltcedar, and the native cottonwood has been completely eliminated. In recent year, various remediation and restoration efforts (e.g., mechanical, chemical, and biological efforts) have been made by local, state, and federal governments to control saltcedar. The distribution map of saltcedar and stratified estimate of invasion area are thus critical for governments to investigate the ecological role of saltcedar in riparian ecosystems and develop corresponding cost-effective control strategies.

4.3. Pixel purity in the MSAC model

Additionally, the composite image was also evaluated according to the purity level of saltcedar within a pixel. The classification accuracies of the composite image for multiple pixel purity levels (i.e., 50% to 90% with an increment of 5%) were calculated using the testing sample (Fig. 6). When saltcedar fractions changed from 50% to 90%, the stratified estimates of overall accuracy increased from 81.26% to 91.45%. The user's accuracy of saltcedar increased from 83.50% to 92.11%, and the producer's accuracy of saltcedar rose from 70.14% to 83.72%. This dramatic increase in the classification accuracies (especially the producer's accuracy of saltcedar) was mostly attributable to two reasons. First, the MSAC model organizes the spectral signatures in a time series to estimate critical phenological transition dates (e.g., time series spectral clustering and time series spectral matching). As the pixel becomes purer, the temporal pattern of spectral signatures along the time series is more pronounced and the noticeable change in the spectral signature at the phenological transition date is more likely to be detected. For each individual pixel, the image selected to construct the composite image is more inclined to be acquired during the leaf senescence stage of saltcedar. Second, when the fractional coverage of saltcedar within a pixel increases, the within-class spectral variability will decrease. The omission and commission errors caused by the spectral confusion between classes are more likely to be corrected.

As the fractional coverage of saltcedar in a pixel affects the shape of its spectral signature and the MSAC model is constructed based on time series of spectral signatures, temporal pattern of spectral signatures along the time series may not be obvious in a mixed pixel. Thus the purity level of the pixel can affect the efficacy of the proposed model in estimating saltcedar leaf senescence timing. Despite the increasing mixed level within a pixel, the performance of the composite image (e.g., pixel purity level being 50%) is still much greater than that of a single Landsat image (see Section 4.2). Yet to conduct the detailed saltcedar mapping, acquiring fine resolution imagery during the key phenological stage would be beneficial for further reconnaissance of saltcedar distributions.

5. Discussion

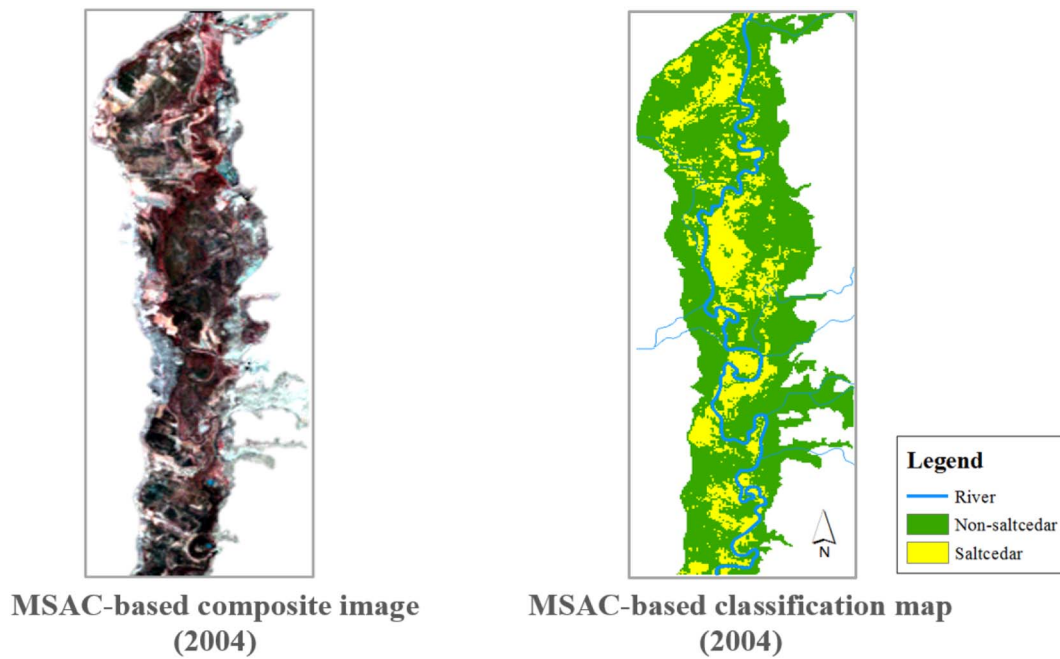Among all the earth observation satellites, Landsat represents the

Table 4

Confusion matrices constructed for the MSAC-based and NDVI-based composite Landsat images.

MSAC-based composite image							NDVI-based composite image					
		Reference data							Reference data			
		Saltcedar	Non-saltcedar	Total	Map area (ha)				Saltcedar	Non-saltcedar	Total	Map area (ha)
Classified data	Saltcedar	334	66	400	946		Classified data	Saltcedar	316	84	400	914
	Non-saltcedar	80	320	400	1649			Non-saltcedar	117	283	400	1681
	Total	414	386	800	2595			Total	433	367	800	2595
	Overall accuracy = 81.26%							Overall accuracy = 73.66%				
	Saltcedar producer's accuracy = 70.14%							Saltcedar producer's accuracy = 59.49%				
	Saltcedar user's accuracy = 83.50%							Saltcedar user's accuracy = 79.00%				

Note: The accuracies of MSAC-based and NDVI-based composite images were evaluated using stratified estimators.

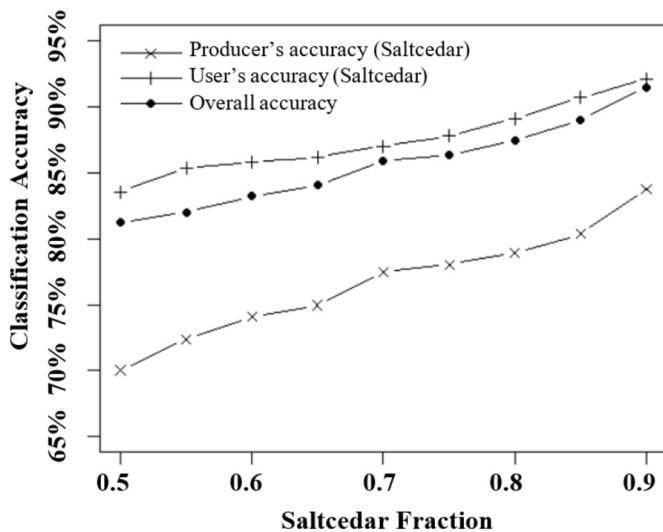**Fig. 5.** Classification map of the MSAC-based composite image in 2004.

Fig. 6. Influence of pixel purity on the classification accuracy of the MSAC-based composite image.

suitable tradeoff between resolution, temporal span, and frequency of coverage, to distinguish the plant species and map their distributions over space and time. Time series of Landsat imagery provides an effective way to estimate plant phenological dynamics in response to climate change and variability. Yet due to the limited temporal resolution, it is often challenging to estimate phenological characteristics using conventional methods. Based on spectral similarities and temporal autocorrelation during fall senescence processes, spectral signatures acquired across multiple years were leveraged to temporally predict the fall phenology of plant species in image-lacking years. The phenological transition date estimated from the MSAC model helped locate the optimal image data to be used for saltcedar monitoring for each individual pixel. Hence this phenology-guided composite image could accommodate the spatial and inter-annual variation in plant phenology, and achieved greater classification accuracy than any single Landsat image (e.g., acquired on Dec. 28 or Nov. 29) in 2004. It is acknowledged that the accuracy assessment based on our reference data

(resampled AISA classification image) was somewhat optimistic. The literature provides extensive guidance on some important issues regarding reference data such as classification errors, residual variability, and spatial autocorrelation, which are beyond the scope of this current study and need to be investigated in future studies (Condés and McRoberts, 2017; Foody, 2010; McRoberts et al., 2016). With the synthesized Landsat time series, the MSAC model provides a new phenological perspective to construct the composite image, which can facilitate the repetitive mapping of invasive saltcedar over wide geographical regions.

The MSAC model provides a new solution to estimate plant phenological dynamics with multiyear time series of spectral signatures. Compared to time series of vegetation index, time series of spectral signatures carries richer information of plant properties (e.g., leaf pigment content, cell structure, and water content). To date, more than 150 vegetation indices (e.g., NDVI, EVI, soil-adjusted vegetation index [SAVI], and NDWI) have been developed in scientific literature, with each accentuating a particular plant property. Different vegetation indices can affect estimation results of phenological transition dates notably. Tracking the collective change of spectral signature along the temporal trajectory may be a desirable alternative. The study demonstrated that the phenology-guided composite image constructed by time series of spectral signatures achieved a more accurate saltcedar mapping result than the NDVI-based composite image. The overall accuracy increased from 73.66% (NDVI-based) to 81.26% (MSAC-based).

Further, we found that time series of spectral signatures were more comparable both within and between years. Unlike vegetation index, spectral signatures acquired across years can be clustered and organized based on their spectral similarity (e.g., shape similarity or cosine distance). Compared to one-dimensional vegetation index, multi-dimensional spectral signatures along the time series exhibit higher temporal autocorrelation. Time series of spectral signatures acquired across multiple years can be more easily integrated to conduct the temporal interpolation, which yields richer phenological trajectory information. Despite being developed for estimating saltcedar phenological transition dates, the components proposed in the MSAC model (i.e., time series spectral outlier removal, time series spectral clustering, and time series spectral matching) could set up a general workflow for exploring plant phenological dynamics given limited Landsat imagery. The MSAC

model may also be generalized to other plant species (e.g., deciduous tree species and agricultural crops) for understanding their phenology. Yet we found that the performance of the MSAC model depends on the pixel purity level. As a pixel becomes more mixed (e.g., more heterogeneous areas), it is more difficult to predict the temporal pattern of spectral signatures along the time series, and to estimate the phenological transition dates.

In this study, the MSAC model utilizes three years' spectral signatures to temporally predict the plant phenology in a certain year. It assumes that saltcedar fractions remain relatively stable during these three years, so as to conduct the temporal interpolation and construct the composite image. This assumption was made according to our extensive field work in the study area, as well as similar climate and hydrological conditions during this period. Yet the robustness of the MSAC model to the violation of the assumption needs to be tested in highly dynamic invasive periods, which can further shed light on how the length of image acquisition periods affects the efficacy of synthesized Landsat time series. In some cases, if saltcedar fractions do not change much over a longer period of time, the corresponding Landsat imagery can be leveraged together to construct the MSAC model. With the increasing possibility of images acquired during the leaf senescence stage being selected, the MSAC model would be more capable of monitoring saltcedar distributions. To overcome the challenge of the coarse temporal resolution of Landsat, in future studies it may be beneficial to construct the time series by fusing the satellite imagery of fine temporal resolution (e.g., MODIS) with Landsat imagery using data fusion algorithms (e.g., Spatial and Temporal Adaptive Reflectance Fusion Model) (Gao et al., 2006). It may also be desirable to incorporate recently available Sentinel-2 and near-surface remote sensing (e.g., PhenoCam) data in constructing the synthesized time series.

The phenological transition date estimated from the MSAC model was used to guide the construction of the composite image. Due to limited temporal resolution of Landsat imagery, the phenological transition date may not directly correspond to the onset of leaf senescence. In this study it denotes that the corresponding image is acquired during the leaf senescence stage. In fact, remote monitoring of phenological events is a subjective process (White et al., 1997). A variety of methods have been developed to estimate the onset of leaf senescence from satellite observations, each suited to a specific research question. Yet there is no consistent definition of leaf senescence from the satellite perspective. Through leveraging the time series of spectral signatures, the MSAC model developed in the study offers new insights into the remote monitoring of vegetation phenological processes. As changes in vegetation phenology show strong couplings to climate and net ecosystem productivity, the MSAC model can also help enhance the understanding of the effects of climate change on ecosystem functioning and biosphere-atmosphere interactions.

As the third most dominant woody plant along riparian corridors in the western US, saltcedar has spread remarkably fast during the past century. The rapid expansion of invasive saltcedar has considerably altered the hydrological regimes and threatened native biotic communities. To effectively control this species and conduct the systemic restoration of riparian ecosystems, it is crucially essential to monitor the distribution of saltcedar and its spatio-temporal dynamics over wide geographical regions. Yet the repetitive monitoring of saltcedar distribution at the regional scale presents many challenges and limitations, largely due to the spatial and inter-annual variation in saltcedar phenology. To accommodate the variation, the MSAC model was developed in this study to construct the phenology-guided composite image using the synthesized Landsat time series. With the readily available historical archives and continuing acquisitions of Landsat imagery, the proposed method shows great promise in facilitating the region-wide monitoring of saltcedar and reevaluating the ecological effects of the species on riparian ecosystems.

6. Conclusions

Over the last century, the rapid expansion of saltcedar along riparian corridors has gravely altered the landscape structures and disrupted the ecosystem functioning in the western US. The MSAC model was developed in this study as a workflow to monitor the inter-annual leaf senescence of saltcedar with limited Landsat imagery. The MSAC model integrates the spectral signatures across years to generate the synthesized time series for estimating phenological transition dates. By accommodating spatial and inter-annual variations in plant leaf senescence timing, the MSAC-based composite image mapped saltcedar distribution more accurately than any single Landsat image in 2004. It also achieved greater accuracy than the monitoring strategy of conventional NDVI-based composite image. Yet built upon the time series of spectral signatures, the performance of the MSAC model may be affected by the pixel purity level, as mixed pixels can obscure the temporal signature patterns along the leaf phenological trajectory. Through characterizing the leaf senescence process, the MSAC model provides unique opportunities to guide the construction of representative composite Landsat images for repetitive mapping of saltcedar distributions over wide geographical regions.

References

Arthur, D., Vassilvitskii, S., 2007. k-Means + +: the advantages of careful seeding. In: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. Society for Industrial and Applied Mathematics, New Orleans, Louisiana, pp. 1027–1035.

Beck, P.S.A., Atzberger, C., Høgda, K.A., Johansen, B., Skidmore, A.K., 2006. Improved monitoring of vegetation dynamics at very high latitudes: a new method using MODIS NDVI. Remote Sens. Environ. 100, 321–334.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32.

Condés, S., McRoberts, R.E., 2017. Updating national forest inventory estimates of growing stock volume using hybrid inference. For. Ecol. Manag. 400, 48–57.

Di Tomaso, J.M., 1998. Impact, biology, and ecology of saltcedar (*Tamarix* spp.) in the Southwestern United States. Weed Technol. 12, 326–336.

Diao, C., Wang, L., 2014. Development of an invasive species distribution model with fine-resolution remote sensing. Int. J. Appl. Earth Obs. Geoinf. 30, 65–75.

Diao, C., Wang, L., 2016a. Incorporating plant phenological trajectory in exotic saltcedar detection with monthly time series of Landsat imagery. Remote Sens. Environ. 182, 60–71.

Diao, C., Wang, L., 2016b. Temporal partial unmixing of exotic salt cedar using Landsat time series. Remote Sensing Letters 7, 466–475.

Dudley, T.L., Deloach, C.J., 2004. Saltcedar (*Tamarix* spp.), endangered species, and biological weed control—can they mix? Weed Technol. 18, 1542–1551.

Evangelista, H.P., Stohlgren, J.T., Morisette, T.J., Kumar, S., 2009. Mapping invasive tamarisk (*Tamarix*): a comparison of single-scene and time-series analyses of remotely sensed data. Remote Sens. 1.

Everitt, J.H., Deloach, C.J., 1990. Remote sensing of Chinese tamarisk (*Tamarix chinensis*) and associated vegetation. Weed Sci. 38, 273–278.

Fisher, J.I., Mustard, J.F., Vadeboncoeur, M.A., 2006. Green leaf phenology at Landsat resolution: scaling from the field to the satellite. Remote Sens. Environ. 100, 265–279.

Foody, G.M., 2010. Assessing the accuracy of land cover change with imperfect ground reference data. Remote Sens. Environ. 114, 2271–2285.

Friedman, J.M., Auble, G.T., Shafroth, P.B., Scott, M.L., Merigliano, M.F., Freehling, M.D., Griffin, E.R., 2005. Dominance of non-native riparian trees in western USA. Biol. Invasions 7, 747–751.

Friedman, J.M., Roelle, J.E., Cade, B.S., 2011. Genetic and environmental influences on leaf phenology and cold hardiness of native and introduced riparian trees. Int. J. Biometeorol. 55, 775–787.

Gao, F., Masek, J., Schwaller, M., Hall, F., 2006. On the blending of the Landsat and MODIS surface reflectance: predicting daily Landsat surface reflectance. IEEE Trans. Geosci. Remote Sens. 44, 2207–2218.

Glenn, E.P., Nagler, P.L., 2005. Comparative ecophysiology of *Tamarix ramosissima* and native trees in western U.S. riparian zones. J. Arid Environ. 61, 419–446.

Graf, W.L., 1978. Fluvial adjustments to the spread of tamarisk in the Colorado plateau region. Geol. Soc. Am. Bull. 89, 1491–1501.

Hamada, Y., Stow, D.A., Coulter, L.L., Jafolla, J.C., Hendricks, L.W., 2007. Detecting tamarisk species (*Tamarix* spp.) in riparian habitats of Southern California using high spatial resolution hyperspectral imagery. Remote Sens. Environ. 109, 237–248.

Hartigan, J.A., 1975. Clustering Algorithms. Wiley, New York.

Hufkens, K., Friedl, M., Sonnentag, O., Braswell, B.H., Milliman, T., Richardson, A.D., 2012. Linking near-surface and satellite remote sensing measurements of deciduous broadleaf forest phenology. Remote Sens. Environ. 117, 307–321.

Hultine, K.R., Belnap, J., van Riper, C., Ehleringer, J.R., Dennison, P.E., Lee, M.E., Nagler, P.L., Snyder, K.A., Uselman, S.M., West, J.B., 2010. Tamarisk biocontrol in the western United States: ecological and societal implications. Front. Ecol. Environ. 8,

467–474.

Ji, W., Wang, L., 2016. Phenology-guided saltcedar (*Tamarix* spp.) mapping using Landsat TM images in western U.S. Remote Sens. Environ. 173, 29–38.

Jonsson, P., Eklundh, L., 2002. Seasonality extraction by function fitting to time-series of satellite sensor data. IEEE Trans. Geosci. Remote Sens. 40, 1824–1832.

Kovalskyy, V., Roy, D.P., Zhang, X.Y., Ju, J., 2012. The suitability of multi-temporal web-enabled Landsat data NDVI for phenological monitoring — a comparison with flux tower and MODIS NDVI. Remote Sensing Letters 3, 325–334.

Kriegel, H.-P., Hubert, M.S., Zimek, A., 2008. Angle-based outlier detection in high-dimensional data. In: Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, Las Vegas, Nevada, USA, pp. 444–452.

Kruse, F.A., Lefkoff, A.B., Boardman, J.W., Heidebrecht, K.B., Shapiro, A.T., Barloon, P.J., Goetz, A.F.H., 1993. The spectral image processing system (SIPS)—interactive visualization and analysis of imaging spectrometer data. Remote Sens. Environ. 44, 145–163.

Lloyd, D., 1990. A phenological classification of terrestrial vegetation cover using shortwave vegetation index imagery. Int. J. Remote Sens. 11, 2269–2279.

MacQueen, J., 1967. Some methods for classification and analysis of multivariate observations. In: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, pp. 281–297 (Oakland, CA, USA).

Masek, J.G., Vermote, E.F., Saleous, N.E., Wolfe, R., Hall, F.G., Huemmrich, K.F., Feng, G., Kutler, J., Teng-Kui, L., 2006. A Landsat surface reflectance dataset for North America, 1990–2000. IEEE Geosci. Remote Sens. Lett. 3, 68–72.

McRoberts, R.E., Chen, Q., Domke, G.M., Ståhl, G., Saarela, S., Westfall, J.A., 2016. Hybrid estimators for mean aboveground carbon per unit area. For. Ecol. Manag. 378, 44–56.

Morisette, J.T., Jarnevich, C.S., Ullah, A., Cai, W., Pedelty, J.A., Gentle, J.E., Stohlgren, T.J., Schnase, J.L., 2006. A tamarisk habitat suitability map for the continental United States. Front. Ecol. Environ. 4, 11–17.

Morisette, J.T., Richardson, A.D., Knapp, A.K., Fisher, J.I., Graham, E.A., Abatzoglou, J., Wilson, B.E., Breshears, D.D., Henebry, G.M., Hanes, J.M., Liang, L., 2009. Tracking the rhythm of the seasons in the face of global change: phenological research in the 21st century. Front. Ecol. Environ. 7, 253–260.

Moulin, S., Kergoat, L., Viovy, N., Dedieu, G., 1997. Global-scale assessment of vegetation phenology using NOAA/AVHRR satellite measurements. J. Clim. 10, 1154–1170.

Nagler, P.L., Glenn, E.P., Didan, K., Osterberg, J., Jordan, F., Cunningham, J., 2008. Wide-area estimates of stand structure and water use of *Tamarix* spp. on the lower Colorado River: implications for restoration and water management projects. Restor. Ecol. 16, 136–145.

Nagler, P.L., Glenn, E.P., Jarnevich, C.S., Shafroth, P.B., 2011. Distribution and abundance of saltcedar and Russian olive in the Western United States. Crit. Rev. Plant Sci. 30, 508–523.

Narumalani, S., Mishra, D.R., Burkholder, J., Merani, P.B.T., Willson, G., 2006. A comparative evaluation of ISODATA and spectral angle mapping for the detection of saltcedar using airborne hyperspectral imagery. Geocarto International 21, 59–66.

Olofsson, P., Foody, G.M., Herold, M., Stehman, S.V., Woodcock, C.E., Wulder, M.A., 2014. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 148, 42–57.

Olofsson, P., Foody, G.M., Stehman, S.V., Woodcock, C.E., 2013. Making better use of accuracy data in land change studies: estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sens. Environ. 129, 122–131.

Pham, N., Pagh, R., 2012. A near-linear time approximation algorithm for angle-based outlier detection in high-dimensional data. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM, Beijing, China, pp. 877–885.

Pu, R., Gong, P., Tian, Y., Miao, X., Carruthers, R.I., Anderson, G.L., 2008. Using classification and NDVI differencing methods for monitoring sparse vegetation coverage: a case study of saltcedar in Nevada, USA. Int. J. Remote Sens. 29, 3987–4011.

Reed, B.C., Brown, J.F., VanderZee, D., Loveland, T.R., Merchant, J.W., Ohlen, D.O., 1994. Measuring phenological variability from satellite imagery. J. Veg. Sci. 5, 703–714.

Robinson, T.W., 1965. Introduction, spread, and areal extent of saltcedar (*Tamarix*) in the Western States. In: Professional Paper.

Silván-Cárdenas, J.L., Wang, L., 2010. Retrieval of subpixel *Tamarix* canopy cover from Landsat data along the Forgotten River using linear and nonlinear spectral mixture models. Remote Sens. Environ. 114, 1777–1790.

Stromberg, J.C., Chew, M.K., Nagler, P.L., Glenn, E.P., 2009. Changing perceptions of change: the role of scientists in *Tamarix* and river management. Restor. Ecol. 17, 177–186.

Van Riper, C., Paxton, K.L., O'Brien, C., Shafroth, P.B., McGrath, L.J., 2008. Rethinking avian response to *Tamarix* on the lower Colorado River: a threshold hypothesis. Restor. Ecol. 16, 155–167.

Vitousek, P.M., D'Antonio, C.M., Loope, L.L., Westbrooks, R., 1996. Biological invasions as global environmental change. Am. Sci. 84, 468–478.

Wang, L., Silván-Cárdenas, J.L., Yang, J., Frazier, A.E., 2013. Invasive saltcedar (*Tamarisk* spp.) distribution mapping using multiresolution remote sensing imagery. Prof. Geogr. 65, 1–15.

Wang, L., Zhang, S., 2014. Incorporation of texture information in a SVM method for classifying salt cedar in western China. Remote Sensing Letters 5, 501–510.

White, M.A., Thornton, P.E., Running, S.W., 1997. A continental phenology model for monitoring vegetation responses to interannual climatic variability. Glob. Biogeochem. Cycles 11, 217–234.

Yang, C., Everitt, J.H., Fletcher, R.S., 2013. Evaluating airborne hyperspectral imagery for mapping saltcedar infestations in west Texas. J. Appl. Remote. Sens. 7, 073556.

Ye, H., Kitagawa, H., Xiao, J., 2014. Continuous angle-based outlier detection on high-dimensional data streams. In: Proceedings of the 19th International Database Engineering & Applications Symposium. ACM, Yokohama, Japan, pp. 162–167.

Zavaleta, E., 2000. Valuing ecosystem services lost to *Tamarix* invasion in the United States. In: Invasive Species in a Changing World, pp. 261–300.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., 2004. Climate controls on vegetation phenological patterns in northern mid- and high latitudes inferred from MODIS data. Glob. Chang. Biol. 10, 1133–1145.

Zhang, X., Friedl, M.A., Schaaf, C.B., Strahler, A.H., Hodges, J.C.F., Gao, F., Reed, B.C., Huete, A., 2003. Monitoring vegetation phenology using MODIS. Remote Sens. Environ. 84, 471–475.

Zhang, X., Goldberg, M.D., 2011. Monitoring fall foliage coloration dynamics using time-series satellite data. Remote Sens. Environ. 115, 382–391.

Zhong, L., Gong, P., Biging, G.S., 2014. Efficient corn and soybean mapping with temporal extendability: a multi-year experiment using Landsat imagery. Remote Sens. Environ. 140, 1–13.