

ORIGINAL ARTICLE

Open Access



Remote sensing of burned areas via PCA, Part 1; centering, scaling and EVD vs SVD

Nikos Alexandris^{1*}, Sandeep Gupta² and Nikos Koutsias³

Abstract

Background: Principal components analysis (PCA) is based conventionally on the eigenvector decomposition (EVD). Mean-centering the input data prior to the eigenanalysis is treated as an integral part of the algorithm. It ensures that the first principal component is proportional to the maximum variance of the input data. Equivalent to EVD, but numerically more robust, is the singular value decomposition (SVD). Mean-centered data subjected to SVD, yield transformation coefficients identical to EVD. Nevertheless, mean-centering is optional in SVD. Avoiding to center the input data, results in generic first component that mainly reflects their mean. This may, however, detect more accurately distinct clusters in PCA-based change detection applications.

Methods: In remote sensing, PCA transforms multi-spectral bands into a new coordinate system. The first, among the transformed components, contain the variance of unchanged landscape features. Succeeding components may contain an enhanced variance of changed features. Such is the case of burned surfaces appearing as distinct clusters in multitemporal composites.

Conclusions: Within this framework, a non-centered SVD may increase the spectral separability of burned clusters among other features in some of the higher order principal components.

Keywords: PCA, EVD, SVD, Mean-centering, Scaling, Burned areas, Spectral enhancement

Background

The goal of this paper is to set the theoretical framework of PCA in remote sensing of burned areas. Specific objectives are to introduce the reader to the (i) EVD-based PCA and the significance of mean-centering and scaling as pre-processing steps; (ii) SVD, an alternative solution to PCA and its difference against EVD; (iii) existing EVD- and SVD-based PCA applications in remote sensing; (iv) the methodological concepts of remotely sensing burned areas via the EVD-based PCA; (v) four SVD-based PCA versions for burned area mapping (vi) the implications of mean-centering and scaling multi-dimensional data prior to PCA. Finally we link the presented theoretical concepts to the remote sensing of burned areas and suggest that a non-centered SVD may perform better than the EVD-based PCA in capturing burned areas.

Principal components analysis

Principal components analysis (PCA) or *transformation* (PCT) is a non-parametric, orthogonal linear transformation of correlated variables [1, 2]. Being probably the oldest and most well-known multivariate analysis technique, PCA is useful in a wide range of applications including data exploration and visualisation of underlying patterns within correlated data sets; decorrelation; detection of outliers; data compression; feature reduction; enhancement of visual interpretability; improvement of statistical discrimination of clusters; ecological ordination; and more [3].

While the transformation can expose the internal structure of multivariate data sets, it is by no means optimised for class separability. PCA does not analyse class labels but uses global statistics to derive the transformation parameters [4]. Hence, there is no guarantee that the directions of maximum variance enhance class separabilities. It is up to the user to identify a high signal-to-noise ratio, via visual or quantitative inspection, of a feature of interest within the principal components. In short, PCA supports cluster

*Correspondence: nik@nikosalexandris.net

¹Independent Researcher, Ebringen, Germany

Full list of author information is available at the end of the article

seeking applications but cannot replace the need for user input.

Applying the EVD-based principal components transformation (henceforth noted as PCA) assumes linearity (or high data correlation); considers the statistical importance of the mean and the variance-covariance; and bases on large variances to highlight useful discrimination properties. In its classical form, the algorithm of the transformation bases on EVD and performs the following steps: **i)** organises the data set in a matrix, **ii)** mean-centers the columns of the matrix (henceforth also referred to as *mean centering* or just *centering*¹), **iii)** calculates the covariance matrix (non-standardised PCA) or the correlation matrix (standardised PCA, a step also known as scaling), **iv)** calculates the eigenvectors and the eigenvalues of the data's variance-covariance or the correlation matrix, **v)** sorts the variances (i.e. the eigenvalues) in decreasing order, and finally **vi)** projects the original dataset signals into what is named Principal Components, or scores, by multiplying them with the eigenvectors which act as weighting coefficients.

Mathematically, PCA can be described as a set of p -dimensional vectors of weights or loadings $w_{(k)} = (w_1, \dots, w_p)_{(k)}$ that map each row vector x_i , of a zero-mean matrix X , to a new vector of principal component scores $t_{(i)} = (t_1, \dots, t_m)_{(i)}$. The scores are given by $t_{k(i)} = x_{(i)} \cdot w_{(k)}$ for $i = 1, \dots, n$ and $k = 1, \dots, m$ in a way that the individual variables of t capture successively the maximum possible variance from x , with each loading vector w constrained to be a unit vector [3]. The complete decomposition of X can be given as

$$Y = W'X \quad (1)$$

where W is a p -by- p matrix (noting that W' is transposed) whose columns are the eigenvectors of $X'X$ determined by the algorithm and Y the matrix of the transformed data. In the case of remotely sensed multispectral images, we consider n spectral response observations (pixel values) on b spectral bands. Following the steps outlined above, the algorithm using the covariance matrix

1. Starts with arranging the spectral responses as n vectors x_1, \dots, x_n and placing them as rows in a single matrix X of dimensions $n \times b$ where the columns correspond to the b spectral bands.
2. For each band b , its mean m is subtracted from all of its spectral response values. Hence all bands present now a zero mean.
3. Next, it calculates the covariance (or correlation) matrix Σ_x of the input data given by

$$\Sigma_X = \frac{\sum_{k=1}^n (x_k - m)(x_k - m)'}{n - 1} \quad (2)$$

We emphasize that calculating the covariance matrix, does actually mean-center the data beforehand. It is of explanatory importance, however, useful to present it as a separate step. The covariance describes the scatter or spread of the spectral responses in the multispectral space. It is symmetric and therefore has orthogonal eigenvectors and real eigenvalues.

Alternatively, to perform PCA on the correlation matrix, instead of the covariance matrix, each band is standardized, i.e. divided by its standard deviation.

4. Successively, the algorithm computes the matrix D of eigenvectors which diagonalises the covariance matrix Σ_x . The respective equation is

$$\Lambda = D' \Sigma_x D \quad (3)$$

where Λ is the diagonal matrix of eigenvalues of Σ_x . It is shown that the covariance matrix of the transformed data Σ_y can be identified as the diagonal matrix of eigenvalues of Σ_x [5]. The relationship between Σ_y and Σ_x is

$$\Sigma_y = D' \Sigma_x D \quad (4)$$

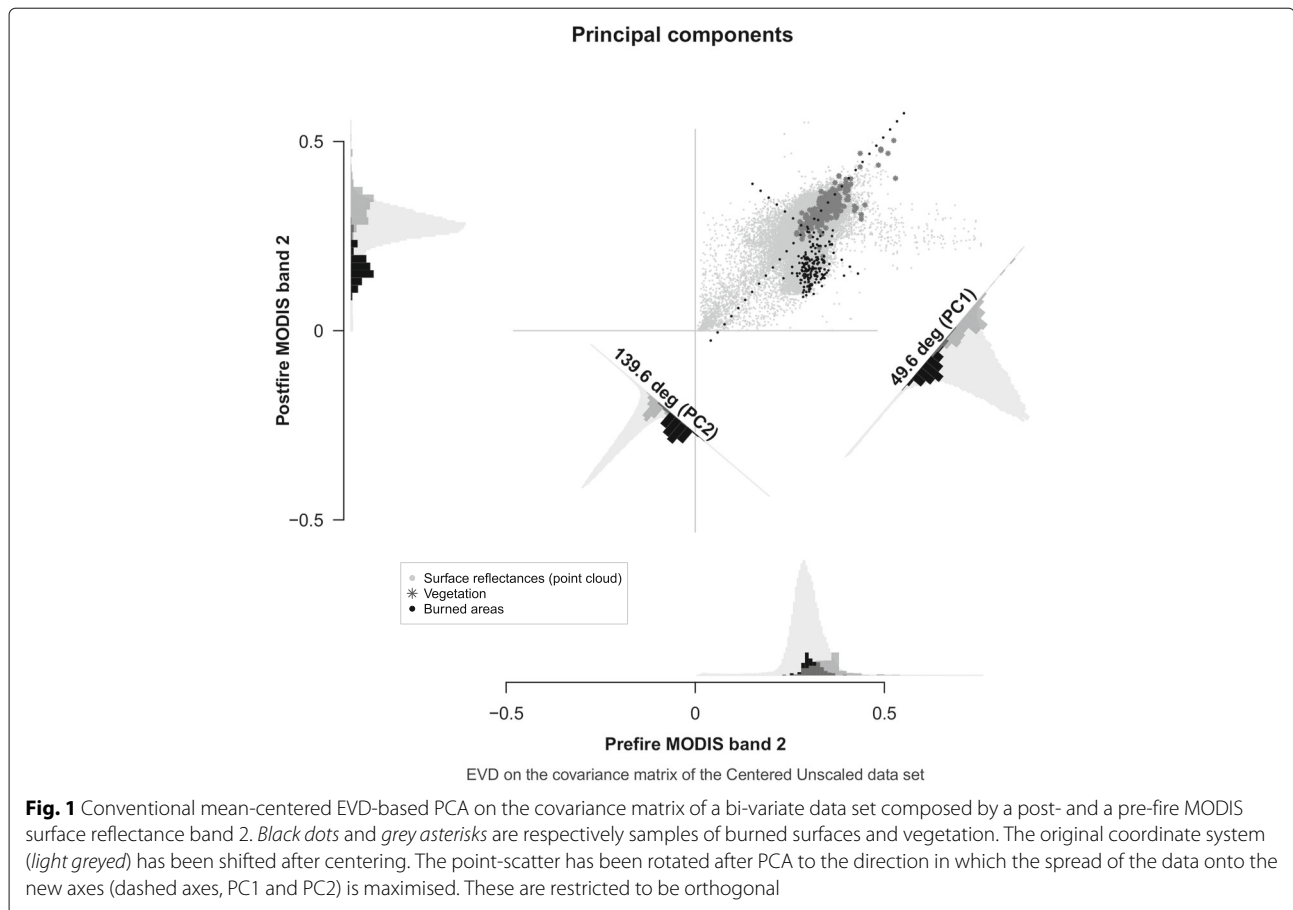
where Σ_y the covariance of the transformed data, which must be diagonal, and D the matrix of eigenvectors of Σ_x , provided that D is an orthogonal matrix.

5. The last step sorts the eigenvectors in D and the eigenvalues in matrix Λ in decreasing order. The eigenvector with the highest eigenvalue, which corresponds to the highest variance, is the first principal component of the data.

Mean-centering, scaling and the eigendecomposition

Mean centering the multivariate data matrix is considered to be integrated in EVD. It is achieved by subtracting the mean of each variable from its own observations. This results in a zero mean vector of observations [6] and successively a zero mean matrix. Graphically, this action translates to an origin shift of the original scatter plot in its gravity center without altering its shape (Fig. 1). This is justified in terms of finding a basis that minimises the mean square error of approximating the original data.

Scaling the centered variables to have unit variance before the analysis takes place is optional. This action, also referred as standardisation, forces the variables to be of equal magnitude by altering the range of the point swarm (Fig. 2). This is required when variables are measured in different units and may be useful when their ranges vary substantially. We note that scaling uncentered variables does not result in unit variance. The usefulness of this combination may even be questionable from a mathematical point of view. However, we include this



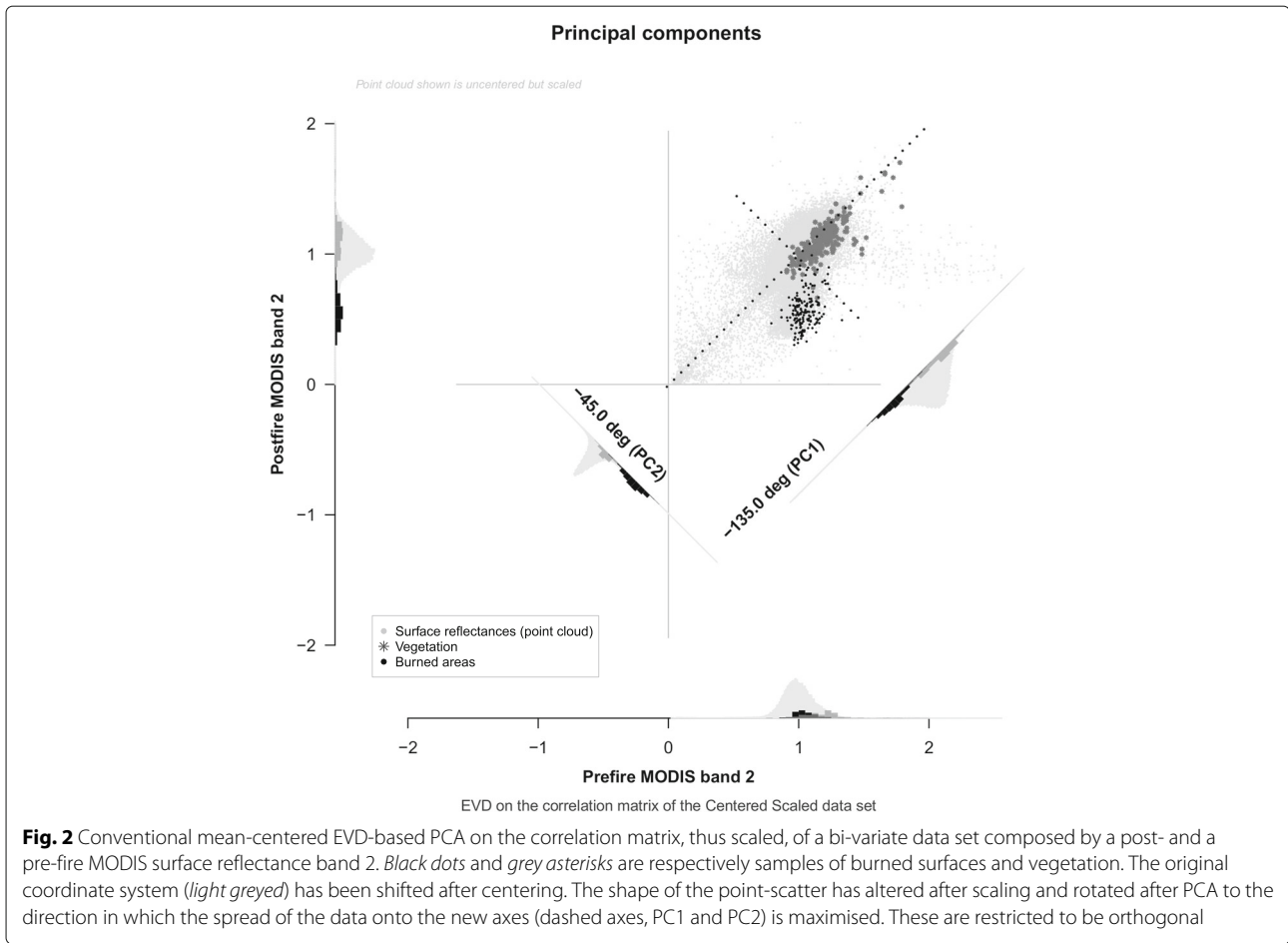
combination in our analytical framework for the sake of experimental completeness.

The *eigendecomposition* of a symmetric matrix yields eigenvectors. Each eigenvector has a positive eigenvalue. The *eigenvectors* define the direction of the variation and the *eigenvalues* define proportionally the length of the axes of variation [7]. Geometrically, the first principal component points in the direction with the largest variance. The second component; being orthogonal to the first, points to the second largest variance. The same pattern is repeated for all remaining components.

PCA rotates actually the point-scatter around its centroid and aligns to the coordinate axes so as to maximise the spread of the data projected onto them. It may be for instructive reasons that authors refer to PCA as *a rotation of the coordinate axes* [8]. This spread is the sum of squares of the data's coordinate points along the axes, also known as *principal component scores*. In addition, the new axes are not correlated. The scores individuated by centered data are expressed as deviations from the mean of the variable. The rotation translates mathematically in a weighted linear combination of the original dimensions.

SVD, an alternative solution to PCA

Another matrix decomposition algorithm which, among other uses, computes the least squares fitting of data, is the Singular Value Decomposition (SVD). The algorithm yields the minimum number of dimensions required to represent a matrix or linear transformation. Clark and Clark [9] describe SVD as being essentially equivalent to a least squares method but numerically more robust according to [10]. Wu [11] notes that better numerical properties are important when higher order dimensions are required. Specifically, when computing the covariance matrix (equation 2), prior to EVD, it is the multiplication of the matrix X by itself that may lead to loss of numerical accuracy. EVD, upon which the conventional PCA is based, and SVD have similar properties and products. Furthermore, PCA can be seen as SVD applied on a column-centered data matrix [6]. Both algorithms are non-sequential and extract hidden variables simultaneously. Yet they differ in several aspects. EVD works on the covariance (or correlation) matrix while SVD operates on the raw data matrix. The main difference of interest, in this work however, is that centering the columns of the input data is performed by default within the framework of EVD-based PCA, while it is optional in SVD.



To describe SVD mathematically, let X denote a matrix of n observations by p variables. The singular value decomposition of X is

$$X = ULA' \quad (5)$$

where U , A are respectively $n \times r$ and $p \times r$ matrices, each of which have orthonormal columns so that $U'U = I_r$ and $A'A = I_r$; L is an $r \times r$ diagonal matrix and r is the rank of X . The columns of U are orthogonal unit vectors of length n called the left singular vectors of X . The columns of A are orthogonal unit vectors of length p called the right singular vectors of X .

Jolliffe [3] states that SVD provides a computationally efficient method of actually finding PCs and that it provides additional insight into what a PCA actually does. He proves the relation between SVD and PCA in that

$$ULA' = X \sum_{k=1}^p a_k a_k' = X \quad (6)$$

where a_k , $k = (r + 1), (r + 2), \dots, p$, are eigenvectors of $X'X$ corresponding to zero eigenvalues. Essentially, the

right singular vectors A of X are equivalent to the eigenvectors of $X'X$, while the singular values in L are equal to the square roots of the eigenvalues of $X'X$. These correspond to the coefficients and standard deviations of the principal components for the covariance matrix.

Applications in remote sensing

Change detection via PCA

In remote sensing studies, PCA is among the most common change detection techniques [12]. The first newly derived components hold the variance related to unchanged landscape features, while succeeding components may feature an enhanced variance of the changed features. Early works demonstrate the usability of PCA in remote sensing ([13–20]). Multi-spectral bands treated as columns build up a data matrix, which is then centered and whose variance-covariance (or correlation) matrix is the basis for the components extraction algorithm. In addition, past literature on the subject, includes various explorations such as using unitemporal [17, 21, 22] and multi-temporal [18, 21, 23–28] multi-spectral data sets. These studies clearly depict multi-temporal approaches as more

effective in terms of both qualitative and quantitative accuracy.

Other studies consider issues related to computational efficiency, disk storage capacities, and more. For instance, spectral contrast mapping—also known as selective PCA—uses two bands that are only weakly correlated to each other as an input and produces two components from which the second summarizes the changes [23, 25, 29]. Harsanyi and Chang [30] developed a method called noise-adjusted PCA in which they suppress undesired spectral signatures in order to maximize the signal-to-noise ratio of a particular signature of interest. The result is a single component identified as a classification image of the spectral feature of interest. Noteworthy in their work, is the fact that centering was not performed before the transformation. Jia and Richards [4] proposed a segmented and possibly multi-staged PCA to deal with computational load issues as well as biases introduced by bands of very high variances. Limitations of this kind have been mostly resolved through the rapid advancement of computer hardware technology allowing researchers to focus more on accuracy issues. For example, [31] demonstrate a multi-temporal and multi-sensor land-use change detection that involves a hybrid classification scheme.

Image processing via SVD

SVD has been extensively used in various disciplines [11, 32]. Its applicability, though, in digital image processing, relies on the fact that multi-spectral remotely sensed spatial data are naturally redundant [33]. Applications include signal estimation in spectral data [34]; classification; data compression and noise reduction [35, 36]; identification of spectral signatures [9]; and feature reduction [33].

To exemplify, [33] compared the performance of SVD to PCA as a feature reduction step on images containing fewer than 100 bands prior to forest/non-forest classifications. In their study, SVD outperformed PCA in terms of classification accuracy and computational time saving. While most SVD-based applications report higher efficacy, in terms of computation and classification results, [9, 37] noted, however, that SVD can be less robust with increasing noise (low signal-to-noise ratio).

Notwithstanding, the use of SVD in remote sensing studies appears somewhat limited. In the past this was attributed to data storage and processing being expensive, especially for very large datasets [34].

Data and concepts

MODIS daily surface reflectance products

The source data for this work are a post- (2007²) and a pre-fire (2006³) MODIS Daily L2G surface reflectance products (MOD09⁴). Their spatial resolution is 250 m and their extent covers Peloponnese in Greece (geographical

extent: south= 36, east= 23.6, north= 38.5, west= 21). These products contain quality filtered and atmospherically corrected observations⁵.

The data consist of infrared observations extracted from MODIS' band 2 (0.858Am). The samples concern burned and vegetated surfaces. These were delineated on-display based on cloud-free and top-quality MODIS observations by utilising MODIS' Quality Data^{6 7 8 9}.

Spectral enhancement of burned area clusters via PCA

PCA can enhance the spectral discrimination between burned and unburned land [16, 21, 23, 24, 38]. The initial poor discrimination of burned observations (pixels) among other classes (such as sparse vegetation and water) is enhanced in some of the higher order principal components. We visualise the rationale behind the technique in bi-dimensional PCA scatterplots (Figs. 1, 2). Regardless, we reiterate that PCA does not isolate distinct clusters—it only enables reviewing data from possibly more useful perspectives.

Figure 1 presents a scatterplot of an untransformed bi-variate data set composed by pre-fire (x-axis) and post-fire (y-axis) surface reflectance observations. The distribution of both the pre- and the post-fire observations are visualised in histograms (light grey coloured) on top of each main axis. Samples of vegetation (dark grey asterisks) and burned areas (black dots) are drawn on top of the surface reflectance observations (grey dots). Their distribution is reflected in mini-histograms (black and dark grey respectively) on top of the larger (grey) histogram in each main axis. The sampled classes overlap entirely onto the x-axis. This is expected as no burn scars exist in the prefire acquisition. However, the samples overlap partially on the y-axis. It is this mixture of spectral information in the postfire acquisition that decreases the separability between the samples in question.

The first step of PCA is to mean-center the input data. While the shape of the scatter remains intact after centering, the coordinate system is shifted to its center. EVD is then applied on the covariance matrix of the mean-centered input data. The transformed data are projected on two new axes. The first principal component (axis rotated by 49.6 degrees, labeled PC1) captures the greatest possible spread of the point cloud. This is the maximum variance of the initial bi-variate data set. The 2nd principal component (axis rotated by 139.6 degrees, labeled PC2) receives a lower amount of data, which corresponds to the second possible maximum variance of the input data.

Likewise as with the input, all transformed observations are represented graphically in offset histograms, yet parallel to the new dimensions. In the first principal component (PC1), the mini-histogram of the burned area samples overlaps partially with the mini-histogram of

Table 1 Four variants of PCA based on SVD

Transformations prior decomposition	Non-scaling	Scaling
Non-centering	uncentered, unscaled (A)	uncentered, scaled (B)
Centering	centered, unscaled (C)	centered, scaled (D)

the vegetation samples. In addition, both mini-histograms are “inside” the main zone of the overall histogram. We observe the similarity of histograms in PC1 with those in the x-axis of the input data. In the second principal component (PC2), however, a higher degree of separation between the sampled classes is obvious. The burned area samples appear as being outside of the main zone of the larger (grey coloured) histogram. In effect, the histograms in PC2 appear similar to those in the y-axis of the input data. Nevertheless, the mini-histograms appear to be slightly more concentrated.

Figure 2 visualises the EVD-based PCA performed on the correlation matrix of the input data. Essentially this means that the input data has been scaled in addition to mean-centering. Comparing to Fig. 1, the centered but unscaled version, a series of important

observations reveal in support to the outlined theoretical concepts. First, the shape of the scatter is altered (more widespread). This is numerically confirmed by the modified range of values $[-2, 2]$ versus $[-0.5, 0.5]$. Second, the mini-histograms of the sampled classes are flattened.

Four ways to extract principal components via SVD

Within the framework of PCA, we question the effects of centering and scaling in burned area clusters. We build data composites based on the following versions of the original data matrix (overviewed in Table 1): (A) uncentered-unscaled, (B) uncentered-scaled, (C) centered-unscaled, (D) centered-scaled. Next, we subject these to the application of SVD-based PCA, aiming to identify improved spectral enhancements of burned area clusters.

Using pre- and post-fire MODIS bands 2, we visualise the application of the four SVD-based PCA versions in Figs. 3, 4, 5 and 6. Their interpretation may follow a path similar as in the previous subsection. Noteworthy is the identity of Figs. 1 and 2 to Figs. 4 and 6 respectively. This is a mere verification that a centered SVD equals the conventional EVD-based PCA.

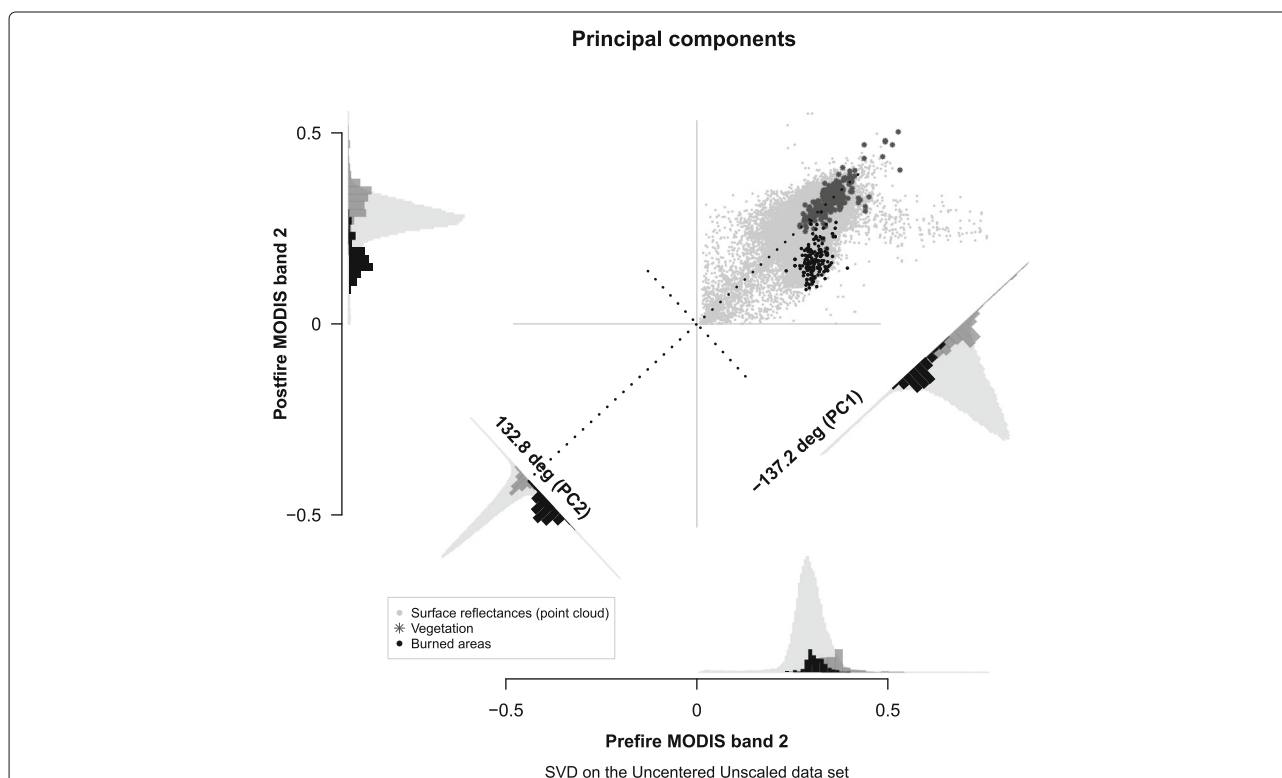
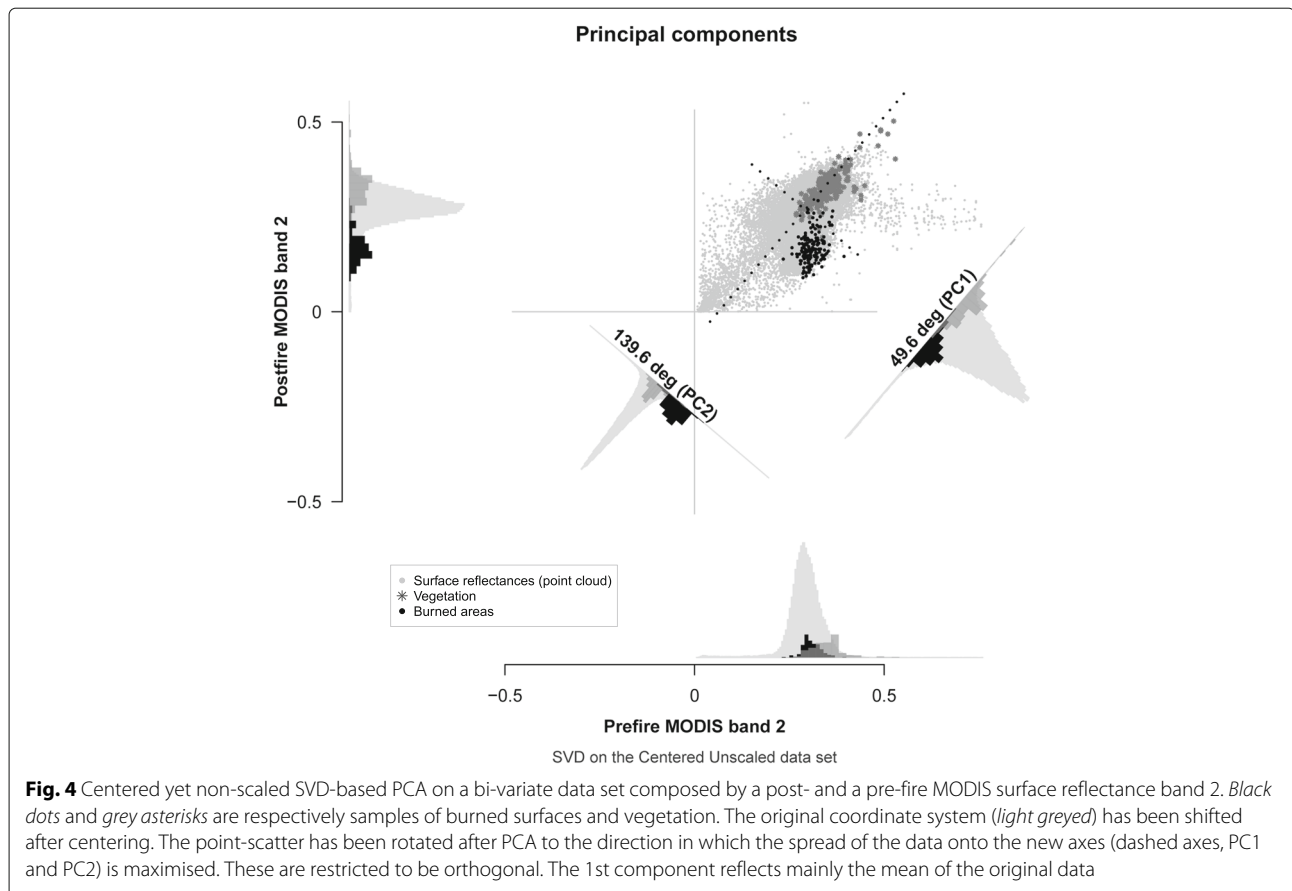


Fig. 3 Non-centered and non-scaled SVD-based PCA on a bi-variate data set composed by a post- and a pre-fire MODIS surface reflectance band 2. Black dots and grey asterisks are respectively samples of burned surfaces and vegetation. The point-scatter has been rotated after PCA to the direction in which the spread of the data onto the new axes (dashed axes, PC1 and PC2) is maximised. These are restricted to be orthogonal. The 1st component reflects mainly the mean of the original data



Most interesting, however, is the non-centered and unscaled SVD case (A) visualised in figure Fig. 3 on page 19. It is unscaled and therefore the input data's range remains unaltered. Non-centering does not shift the origin of the input scatter plot. Careful observation hints to more concentrated mini-histograms when comparing them to their equivalents in the rest of the versions.

Implications of pre-processing data prior to PCA

Effects of mean-centering

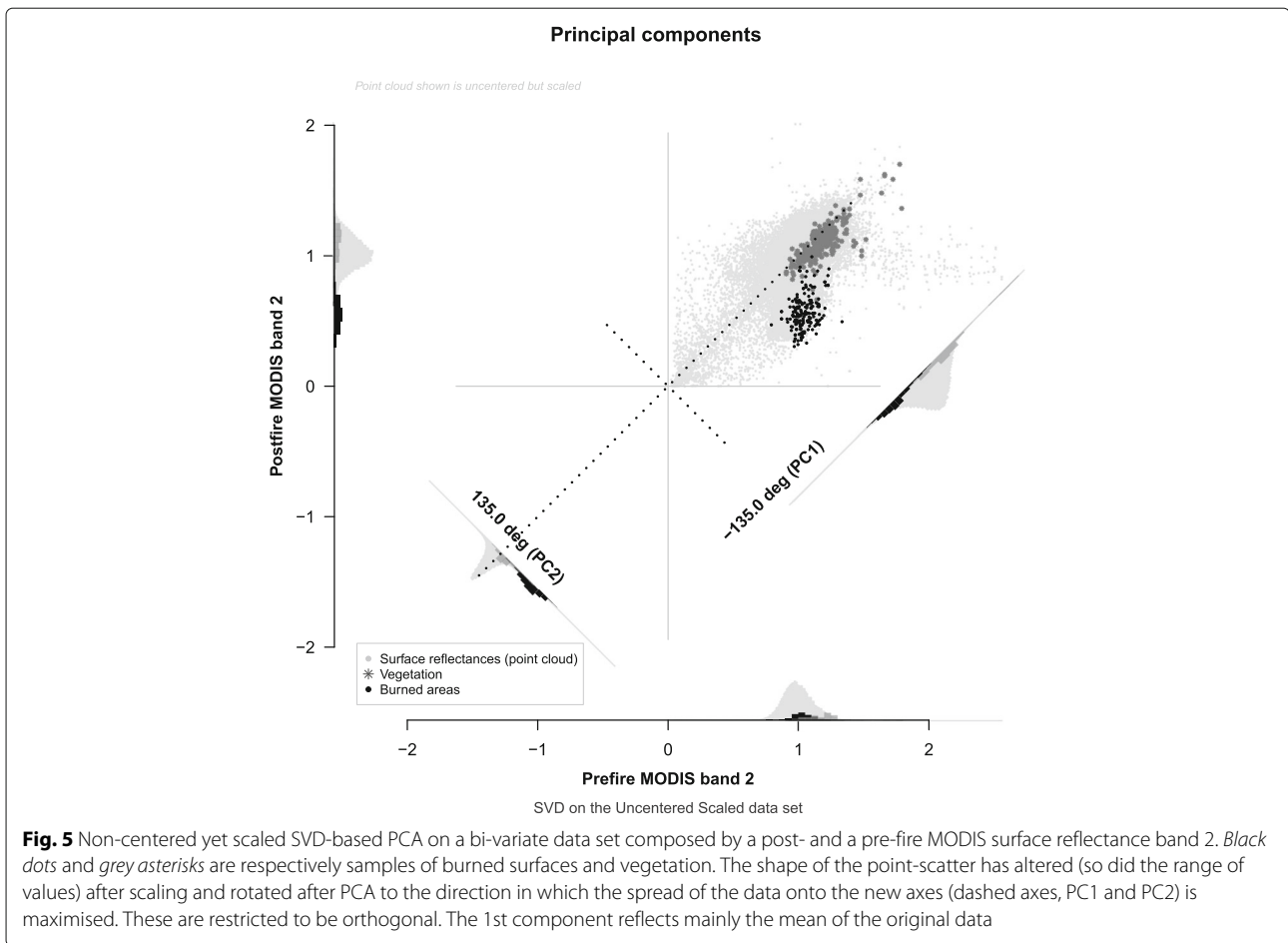
While both PCA and SVD have been widely used in applied scientific research, most studies do not detail why centering is (within the framework of PCA) or should be (within the framework of SVD) performed at all. Few studies in applied disciplines, most of them dating back some decades ago, touch upon this question in-depth ([8, 39–44] or even debate it [45–48] when compared to numerous applications that employ PCA without detailing its mathematical background. In addition, the existence of a recent statistical study shedding light on the issue [6], indicates that the question itself is infamous.

The fact that centering is treated as an integral part of PCA could be a reason why many studies have not

experimented with the differences and effects of applying the general steps of PCA on the raw data matrix. The contrast between a centered and an uncentered PCA lies in the projection (or spread) of the data on to the principal axes. In a centered PCA the spread is proportional to the variance of the data's coordinates, while in an uncentered PCA it is simply the sum of the squared distances of the points from the unshifted origin (non-central second moments). Cadima and Jolliffe [6] explain that *in an uncentred PCA it is variability about the origin, rather than about the centre of gravity of the n-point scatter in \mathbb{R}^p , that will be of concern.*

Non-centering can, and probably will, break the premise of decorrelation between the extracted components. Rather, the first component will reflect the mean instead of the greatest variance [8]. That is, in case the same set of axes are relevant to the explanation of the variability of all clusters (high within-axes heterogeneity), non-centered PCA results in a single generic component [43].

Nonetheless, cases exist where non-centered PCA distinguishes disjunctions more efficiently by intercepting clusters in decreasing order of importance due to size and homogeneity [8, 30, 41, 42]. According to [41], the influence of variance within a cluster on another cluster can



be minimised by not subtracting the mean. Whenever the data form distinct clusters (i.e. clusters have zero, or very small, projections on some subset of the original axes – that is high between-axes heterogeneity), the variability associated with each cluster, is mainly explained by certain axes only. The first few uncentered principal axes pass through (or close to) one of the different clusters, thus *allowing a clearer and simpler picture of the data structure*.

Effects of scaling

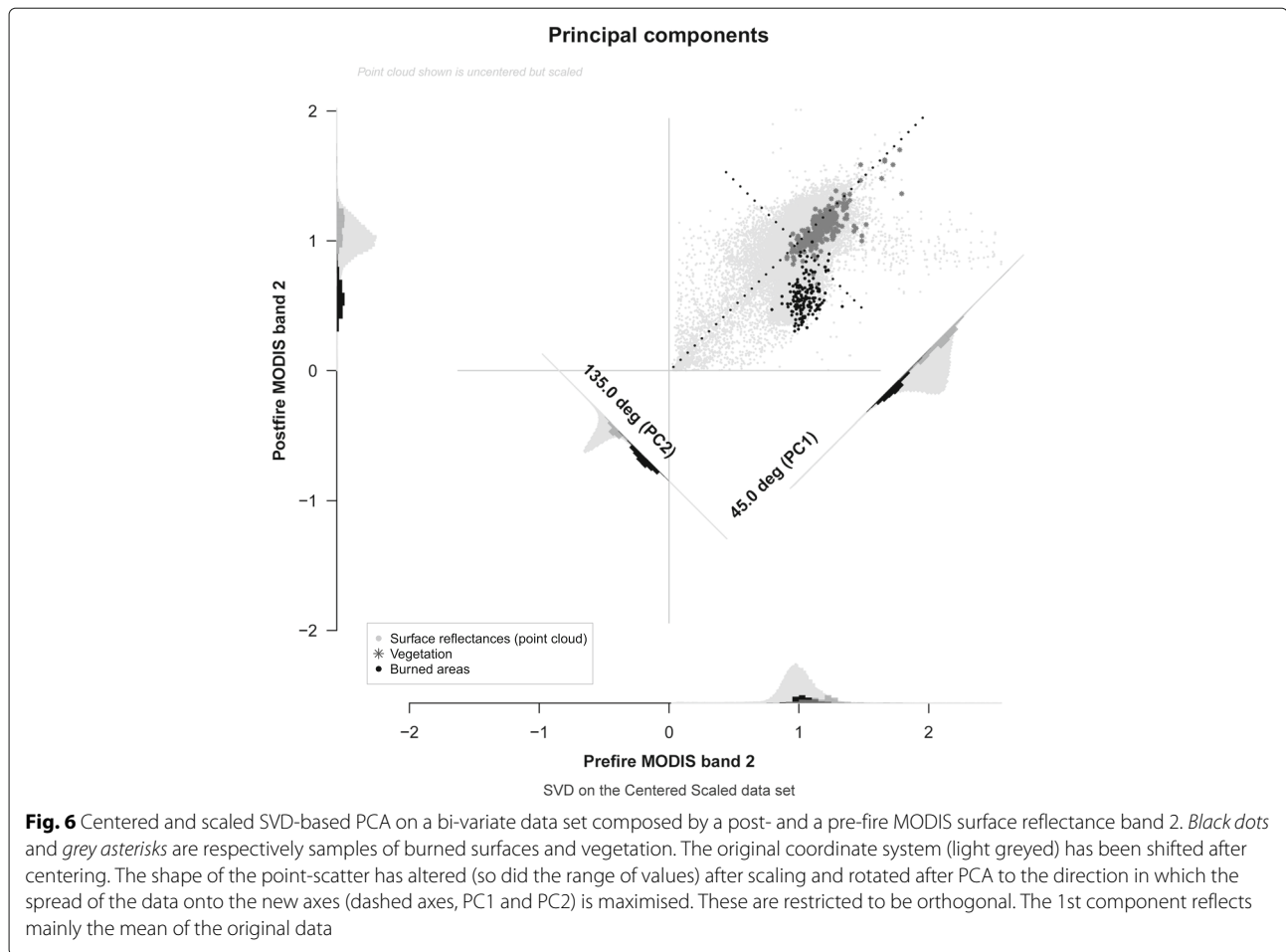
The effects of scaling remotely sensed data, prior to PCA, in terms of subsequent classification accuracy, have been investigated by many including [23, 49–51]. Most of these studies conclude that scaling is useful in preventing certain features from dominating the analysis due to their large numerical values [52, 53]. Even recent applications report greater efficacy by using data scaled to unit variance [31, 54].

Nevertheless, [55] noted that *in the case of using bands with the same physical units, the derivation of the (principal) components from a non-standardised matrix can be justified*. The justification relies on the assumption that

the information contained within each band is a function of the precision with which the spectral data are sensed. In order to exemplify this, [56] justify the use of the variance-covariance matrix, which weights each band according to its relative variance and therefore produces components which are more easy to interpret. Likewise, [57] confirm that standardization produces less noisy high order components and decreases omission errors in the subsequent fire scar classification, but at the cost of increased commission errors. Due to their preference for a conservative fire scar mapping (fewer commission errors), they deliberately use the unstandardised PCA which they find to perform excellent in extracting fire scars.

Linking non-centered SVD to remote sensing of burned areas

The structure of a multi-spectral data set may be statistically retouched, without destroying existing clusters, in numerous ways. Factually, some operations alter the distances between clusters. These may increase or decrease the spectral distances between various land cover classes. Others preserve the existing distances and modify the way



that principal axes intercept existing clusters. The latter is the case for the conventional form of EVD-based PCA. It involves data centering and, optionally, scaling. Thus the data are subject to the implications already discussed.

SVD is a numerically more robust alternative solution to PCA. Hence, it is preferable for high dimensional data sets. Contrary to EVD, centering the raw data matrix in SVD is optional. In case of distinct clusters in high-dimensional data (high between-axes heterogeneity), avoiding centering, minimises the influence of a cluster's variance on another cluster. In addition, the first principal component captures mainly the mean of the input data. This is advantageous since burned area samples are, expectedly, below the mean. In support, uncentered principal axes, pass close to one of the different clusters. Consequently, existing data clusters, may be detected more clearly by a non-centered SVD.

In this context, it is worth experimenting to avoid scaling the data. Subtle yet important variations between clusters may be preserved in the derived components. Such is the case of burned surfaces. They are indexed by lower values due to higher absorption of solar energy in

comparison to other surfaces, especially for a few weeks after the pause of the fires. Because there is little deviation among responses corresponding to the same land category, distinct clusters of data are formed. Burned area clusters feature low internal heterogeneity and are, thus, easier to distinguish in some spectral bands. This is even more pronounced in bi-temporal data sets (composed by pre- and post-fire observations) where the between-axes heterogeneity is even higher.

The theoretical concepts presented in this article are demonstrated numerically via a Multiple Response Procedure [58] and supported by separability metrics in a second article titled *Remote sensing of burned areas via PCA, Part 2: SVD-based PCA using MODIS and Landsat data* [59].

Conclusions

PCA can enhance the spectral separability of burned surfaces among other land cover features. Burned area clusters present low internal heterogeneity and form distinct clusters. In addition, they are easier to detect in some dimensions of a multi-spectral data set. Further

they are absent in some dimensions (i.e. pre-fire data) in multi-temporal composite data set. Therefore, they fit the purpose of a non-centered SVD-based PCA.

Nevertheless, due to the natural redundancy of remotely sensed multi-spectral data, and both EVD and SVD being agnostic about data clusters, it may be complex to interpret the effects of centering and scaling the input data prior to PCA. Thorough examination is necessary to identify optimal data pre-processing options.

Endnotes

¹ Also referred as conversion to *z-scores*, *z-values*, *standard scores*, *normal scores*, *standardised variables*.

² MOD09GQ.A2007242.h19v05.005.2007244231200.

³ MOD09GQK.A2006239.h19v05.004.2006241155630.

⁴ These data are distributed by the Land Processes Distributed Active Archive Center (LP DAAC), located at the U.S. Geological Survey (USGS) Earth Resources Observation and Science (EROS) Center (lpdaac.usgs.gov).

⁵ https://lpdaac.usgs.gov/dataset_discovery/modis/modis_products_table/mod09gq. (Accessed 27 May 2017)

⁶ For MOD09 products visit: http://landweb.nascom.nasa.gov/cgi-bin/QA_WWW/newPage.cgi?fileName=mod09_l2&subdir=v2specs. (Accessed 27 May 2017)

⁷ For Collection4 MODIS products read more at: http://landweb.nascom.nasa.gov/cgi-bin/QA_WWW/detailInfo.cgi?prod_id=MOD09GQK&ver=C4. (Accessed 27 May 2017)

⁸ For Collection5 MODIS products read more at: http://landweb.nascom.nasa.gov/cgi-bin/QA_WWW/detailInfo.cgi?prod_id=MOD09GQ&ver=C5. (Accessed 27 May 2017)

⁹ For Known issues in MOD09 products (surface reflectance): https://landweb.nascom.nasa.gov/cgibin/QA_WWW/getSummary.cgi?esdt=MOD09type=C5. (Accessed 27 May 2017)

Acknowledgments

The authors thank Aniruddha Ghosh and Gregor Trefalt for reading the manuscript.

Authors' contributions

All authors contributed equally to this article. All authors read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Independent Researcher, Ebringen, Germany. ²Institute of Environmental Studies, Kurukshetra University, Kurukshetra, India. ³Department of Environmental and Natural Resources Management, University of Patras, Agrinio, Greece.

Received: 22 December 2016 Accepted: 22 May 2017

Published online: 13 July 2017

References

- Pearson K. On lines and planes of closest fit to systems of points in space. *Philos Mag Series 6*. 1901;2(11):559–72. doi:10.1080/14786440109462720.
- Hottelling H. Analysis of a Complex of Statistical Variables Into Principal Components: Warwick & York; 1933. <http://www.research-bw.de/db/ebsco.php/search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=1934-06132-000&site=ehost-live>.
- Jolliffe IT. *Principal Component Analysis*, 2nd edn: Springer; 2002. 28 illustrations. <http://www.springer.com/statistics/statistical+theory+and+methods/book/978-0-387-95442-4>.
- Jia X, Richards J. Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification. *IEEE Trans Geosci Remote Sensing*. 1999;37(1):538–42.
- Richards J, Jia X. *Remote Sensing Digital Image Analysis. An Introduction*. Third, Revised and Enlarged Edition, 3rd edn: Springer; 1999. p. 363. Hard cover.
- Cadima J, Jolliffe I. On relationships between uncentred and column-centred principal component analysis. *Pak J Stat*. 2009;25(4):473–503.
- Ahearn S, Wee C. Data space volumes and classification optimization of spot and landsat tm data. *Photogrammetric Eng Remote Sensing*. 1991;57(1):61–5.
- Pielou EC. *The Interpretation of Ecological Data: a Primer on Classification and Ordination*, Hardcover edn. New York: Wiley; 1984.
- Clark C, Clark A. Spectral identification by singular value decomposition. *Int J Remote Sensing*. 1998;19(12):2317. doi:10.1080/014311698214749.
- Golub G, Reinsch C. Singular value decomposition and least squares solutions. *Numerische Mathematik*. 1970;14(5):403–20. doi:10.1007/BF02163027.
- Wu W. The kernel pca algorithms for wide data. part i: Theory and algorithms. *Chemometrics Intell Lab Syst*. 1997;36(2):165. doi:10.1016/S0169-7439(97)00010-5.
- Lu D, Mausel P, Brondizio E, Moran E. Change detection techniques. *Int J Remote Sensing*. 2003;25(12):2365. doi:10.1080/0143116031000139863.
- Lodwick GD. Measuring ecological changes in multitemporal Landsat data using principal components. In: *International Symposium on Remote Sensing of Environment*, 13 th, Ann Arbor, Mich; 1979. p. 1131–41.
- Byrne GF, Crapper PF, Mayo KK. Monitoring land-cover change by principal component analysis of multitemporal landsat data. *Remote Sens Environ*. 1980;10(3):175. doi:10.1016/0034-4257(80)90021-8.
- Lodwick GD. A computer system for monitoring environmental change in multitemporal Landsat data. *Can J Remote Sens*. 1981;7:24.
- Richards JA, Milne AK. Mapping fire burns and vegetation regeneration using principal components analysis. In: 1983 *International Geoscience and Remote Sensing Symposium (IGARSS'83)*. San Francisco; 1983.
- Tanaka S, Kimura H, Suga Y. Preparation of a 1:25000 landsat map for assessment of burnt area on etajima island. *Int J Remote Sensing*. 1983;4(1):17. doi:10.1080/01431168308948528.
- Richards JA. Thematic mapping from multitemporal image data using the principal components transformation. *Remote Sensing Environ*. 1984;16(1):35–46. doi:10.1016/0034-4257(84)90025-7.
- Ingebritsen SE, Lyon RJP. Principal components analysis of multitemporal image pairs. *Int J Remote Sensing*. 1985;6(5):687. doi:10.1080/01431168508948491.
- Canas AAD, Barnett ME. The generation and interpretation of false-colour composite principal component images. *Int J Remote Sensing*. 1985;6(6):867. doi:10.1080/01431168508948510.
- Milne A. The use of remote sensing in mapping and monitoring vegetational change associated with bushfire events in eastern australia. *Geocarto Int*. 1986;1(1):25. doi:10.1080/10106048609354022.

22. Patterson M, Yool S. Mapping fire-induced vegetation mortality using landsat thematic mapper data: a comparison of linear transformation techniques. *Remote Sensing Environ.* 1998;65(2):132. doi:10.1016/S0034-4257(98)00018-2.
23. Pereira J. Burned area mapping with conventional and selective principal component analysis. *Finisterra.* 1992;27(53):63.
24. Ribed P, Lopez A. Monitoring burnt areas by principal components analysis of multi-temporal tm data. *Int J Remote Sensing.* 1995;16(9):1577. doi:10.1080/01431169508954497.
25. Siljeström A, Moreno A, Vikgren K, Cáceres L. Technical note the application of selective principal components analysis (spca) to a thematic mapper (tm) image for the recognition of geomorphologic features configuration. *Int J Remote Sensing.* 1997;18(18):3843. doi:10.1080/014311697216658.
26. Rogan J, Yool S. Mapping fire-induced vegetation depletion in the peloncillo mountains, arizona and new mexico. *Int J Remote Sensing.* 2001;22(16):3101. doi:10.1080/01431160152558279.
27. Nielsen T, Mbaw K, Kane R. A statistical methodology for burned area estimation using multitemporal avhrr data. *Int J Remote Sensing.* 2002;23(6):1181. doi:10.1080/01431160110078449.
28. Brewer C, Winne J, Redmond R, Opitz D, Mangrich M. Classifying and mapping wildfire severity: A comparison of methods. *Photogrammetric Eng Remote Sensing.* 2005;71:1311–20.
29. Chavez P, Kwarteng A. Extracting spectral contrast in landsat thematic mapper image data using selective principal component analysis. *Photogrammetric Eng Remote Sensing.* 1989;55(3):339–48.
30. Harsanyi JC, Chang C-I. Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach. *IEEE Trans Geosci Remote Sensing.* 1994;32(4):779–85.
31. Deng J, Wang K, Deng Y, Qi G. Pca-based land-use change detection and analysis using multitemporal and multisensor satellite data. *Int J Remote Sensing.* 2008;29(16):4823. doi:10.1080/01431160801950162.
32. van den Broek W, Wienke D, Melssen W, de Crom C, Buydens L. Identification of plastics among nonplastics in mixed waste by remote sensing near-infrared imaging spectroscopy. 1. image improvement and analysis by singular value decomposition. *Anal Chem.* 1995;67(20):3753. doi:10.1021/ac00116a022.
33. Phillips R, Watson L, Wynne R, Blinn C. Feature reduction using a singular value decomposition for the iterative guided spectral class rejection hybrid classifier. *ISPRS J Photogrammetry Remote Sensing.* 2009;64(1):107–16. doi:10.1016/j.isprsjprs.2008.03.004.
34. Danaher S, O'Mongain E. Singular value decomposition in multispectral radiometry. *Int J Remote Sensing.* 1992;13(9):1771. doi:10.1080/01431169208904226.
35. Herries GM, Danaher S, Murray A. Characterisation of forestry species - a comparison using singular value decomposition (svd) and artificial neural networks (ann). *IEE Conf Publ.* 1995;1995:815–9. doi:10.1049/cp:19950773.
36. Herries G, Selige T, Danaher S. Singular value decomposition in applied remote sensing. *IEE Semin Digests.* 1996;1996(27):5–5. doi:10.1049/ic:19960159.
37. Obiefias W, Soriano M, Saloma C. Svd vs pca: Comparison of performance in an imaging spectrometer. *Sci Diliman.* 2007;16(2):74–8.
38. In: Chuvieco E, editor. *A Review of Remote Sensing Methods for the Study of Large Wildland Fires.* Departamento de Geografía, Universidad de Alcalá, Colegios 2, 28801 Alcalá de Henares, Spain: Printed in Spain; 1997. Megafires project ENV-CT96-0256. <http://www.geogra.uah.es/emilio/pdf/pospaplib.PDF>.
39. Orloci L. Data Centering: A Review and Evaluation with Reference to Component Analysis. *Syst Zool.* 1967;16(3):208–12. <http://www.jstor.org/stable/2412067>.
40. Okamoto M. Four techniques of principal component analysis. *J Jpn Stat Soc.* 1972;2(2):63–9.
41. Noy-Meir I. Data transformations in ecological ordination: I. some advantages of non-centering. *J Ecol.* 1973;61(2):329–41. <http://www.jstor.org/stable/2259029>.
42. Feoli E. On the resolving power of principal component analysis in plant community ordination. *Vegetatio.* 1977;33:119–25. <http://www.jstor.org/stable/20036971>.
43. DiBartolomeo G, Marchetti E. Central Banks and Information Provided to the Private Sector. Roma: Dipartimento di Economia Pubblica, Università di Roma La Sapienza; 2003. Monograph, Working Paper No.61. http://dep.eco.uniroma1.it/docs/working_papers/WP61.pdf.
44. Miranda AA, Borgne YA, Bontempi G. New routes from minimal approximation error to principal components. *Neural Process Lett.* 2008;27(3):197. doi:10.1007/s11063-007-9069-2.
45. Mann ME, Bradley RS, Hughes MK. Reply to mcintyre and mckittrick: Proxy-based temperature reconstructions are robust. *Proc Nat Acad Sci.* 2009;106(6):11. doi:10.1073/pnas.0812936106. <http://www.pnas.org/content/106/6/E11.full.pdf+html>.
46. McIntyre S, McKittrick R. Proxy inconsistency and other problems in millennial paleoclimate reconstructions. *Proc Nat Acad Sci.* 2009;106(6):10. doi:10.1073/pnas.0812509106. <http://www.pnas.org/content/106/6/E10.full.pdf+html>.
47. Wahl ER, Ammann CM. Robustness of the mann, bradley, hughes reconstruction of northern hemisphere surface temperatures: Examination of criticisms based on the nature and processing of proxy climate evidence. *Climatic Change.* 2007;85(1):33–69.
48. Zorita E, González-Rouco F, Legutke S. Testing the Mann et al. (1998) approach to paleoclimate reconstructions in the context of a 1000-yr control simulation with the echo-g coupled climate model. *J Climate.* 2003;16(9):1378–90.
49. Eastman J, Fulk M. Long sequence time-series evaluation using standardized principal components. *Photogrammetric Eng Remote Sensing.* 1993;59:991–6.
50. Guirguis S, Hassan H, El-Raey M, Hussain M. Technical note multi-temporal change of lake brullus, egypt, from 1983 to 1991. *Int J Remote Sensing.* 1996;17(15):2915. doi:10.1080/01431169608949118.
51. Hirose Y. Application of standardized principal component analysis to land-cover characterization using multitemporal avhrr data. *Remote Sensing Environ.* 1996;58(3):267. doi:10.1016/S0034-4257(96)00068-5.
52. Duda R, Hart P. *Pattern Classification and Scene Analysis.* New York, John Wiley-Sons; 1973.
53. Eklundh L, Singh A. A comparative analysis of standardised and unstandardised principal components analysis in remote sensing. *Int J Remote Sensing.* 1993;14(7):1359. doi:10.1080/01431169308953962.
54. Koutsias N, Mallinis G, Karteris M. A forward/backward principal component analysis of landsat-7 etm+ data to enhance the spectral signal of burnt surfaces. *ISPRS J Photogrammetry Remote Sensing.* 2009;64(1):37. doi:10.1016/j.isprsjprs.2008.06.004.
55. Singh A, Harrison A. Standardized principal components. *Int J Remote Sensing.* 1985;6(6):883. doi:10.1080/01431168508948511.
56. García-Haro F, Gilabert M, Melia J. Monitoring fire-affected areas using thematic mapper data. *Int J Remote Sensing.* 2001;22(4):533. doi:10.1080/01431160050505847.
57. Hudak A, Brockett B. Mapping fire scars in a southern african savannah using landsat imagery. *Int J Remote Sensing.* 2004;25(16):3231–243. doi:10.1080/01431160310001632666.
58. Mielke PWJ. The application of multivariate permutation methods based on distance functions in the earth sciences. *Earth Sci Rev.* 1991;31:55–71. doi:10.1016/0012-8252(91)90042-E.
59. Alexandris N, Koutsias N, Gupta S. Remote sensing of burned areas via PCA, Part 2: SVD. based PCA using MODIS and Landsat data. *Open Geospatial Data, Software and Standards.* 2017. doi:10.1186/s40965-017-0029-0.